

QY 2561 IleIysSerValTyrIleuValSerGluThrProIleProThrThrile 2580
 DB 8021 ATCAAGTCGCTGGAGGACTCTCTGGAGACTCGAAGAACCACTATCCCAACCACT 8080
 QY 2581 MetAlaIysAbnGluValPheCysValAspProThrIleGlyIleValAlaAlaArg 2600
 DB 8081 ATGGCCAAATATGAGGTGTCTTGGCTGGACCCCAACAGGGGGGAGAGAGAGCTGCG 8140
 QY 2601 LeuIleValTyrProAspLeuGlyValArgValCysGluIleMetAlaLeuTyrAspIle 2620
 DB 8141 CTTATCGTTTACCTGACTCGGGCTCAGGCTCTGGAGAGAGATGGCCCTTTATGACAT 8200
 QY 2621 ThrGluIleLeuProGluAlaValMetGlyValAspTyrGlyPheGluTyrSerProAla 2640
 DB 8201 ACAAAATCTCTCAGCGGCTGATGGGGCTCTTATGGATTTCAGTATATCCCGGCT 8260
 QY 2641 GluArgValGluPheIleLeuIleValAlaIleGluIleValAspProMetGlyPheSer 2660
 DB 8261 CAGCGGTAGAGTTTCTCTGAAGCATGGCGGAGAAAGAGAGACCTATGGGTATTCG 8320
 QY 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 DB 8321 TATCATACCCGATGCTTTGACTCAACCGTCACTGAGAGAGACATCAGGACTGAGAGTCC 8380
 QY 2681 IleTyrArgAlaCysSerLeuProGluGluAlaIleThrAlaIleHisSerLeuThrGlu 2700
 DB 8381 ATATATGGGCTGCTCTTCCCGAGGAGGCCACACTGCCATACACTGCTCACTGAG 8440
 QY 2701 ArgLeuTyrValGlyIleProMetPheAsnSerIleGlyGluThrCysGlyTyrArgArg 2720
 DB 8441 AGACTTAGCTGGAGGGCTTATGTTCAACGACAGGGCCAACTCTGGGTACAGGCT 8500
 QY 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysGlyValIys 2740
 DB 8501 TGGCGGCCAGCGGGGTCTCACCACATAGATGGGAGACACATACATACATGCTAGTGAA 8560
 QY 2741 AlaLeuAlaCysIleValAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 DB 8561 GCCTTAGCGCTGTAAAGCTGCAGGGATATTCGGCCCACTATGCTGTATGGGGGAT 8620
 QY 2761 AspLeuValIleSerGluSerGluGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 DB 8621 GACTTGGTGTCTATCTCAGAAAGCCGAGGAGCCAGGAGGAGGAGGAGGAGGAGGAG 8680
 QY 2781 PheThrGluAlaMetThrArgTyrSerAlaProProGluAspProArgProGluTyr 2800
 DB 8681 TTCAGCGGCTATGACGAGGTATTCGCTCTCTGGTGACCCCTCCAGACCGGAGTAT 8740
 QY 2801 AspLeuGluIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGluGly 2820
 DB 8741 GATCTGAGCTGATACATCTTGTCTCTCAAAATGTGTCTGGCGCTGGGCGCCAGAGGC 8800
 QY 2821 ArgArgArgTyrIleuThrArgAspProThrProIleAlaArgAlaAlaIleProGlu 2840
 DB 8801 CGCGCAGATATCTACCTGACAGAGACCTACCACTCCAAATCGCCCGGCTGCTGGGAA 8860
 QY 2841 ThrValArgHisSerProValAsnSerIleLeuGluValIleIleGlnTyrAlaProThr 2860
 DB 8861 ACAGTTAGACACTCCCTCTCAATTCATGCTGGAGGAGGAGGAGGAGGAGGAGGAGG 8920
 QY 2861 IleThrAlaArgMetValLeuMetThrIlePhePheSerIleLeuMetAlaGlnAspThr 2880
 DB 8921 ATATGGCTCGCATGGTCTCTGATGACACTCTCTCTCCATCTCTCAAGCTCAAGACAG 8980
 QY 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 DB 8981 CTGACAGAGAACTCACTTTCAGATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9040
 QY 2901 LeuProAlaIleIleGluArgLeuHisGlyIleuAspAlaPheSerLeuHisThrThr 2920
 DB 9041 CTCGAGCTATATTAAGGAGTATCATGGCTTTCAGCTTTTCTCTGACACATACACT 9100
 QY 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgIysLeuGlyAlaProLeuArg 2940

DB 9101 CCCAGCAACTGACAGGGTGGCTTTCAGCCCTCAGAAACTTGGGGCCACCCCTCAGA 9160
 QY 2941 AlaTyrPheSerArgAlaArgAlaValArgAlaSerIleuLeuIleSerArgGlyGlyArgAla 2960
 DB 9161 GCCTGGAAAGAGCCGGGACGTCAGTTCAGGGCTCTCATCTCCGCTGGGGGAGAGCG 9220
 QY 2961 AlaValCysGlyValTyrTyrIleuPheAsnTrpAlaValIleThrIleuLeuIleThrPro 2980
 DB 9221 GCGTITGGGTCGATATCTCTTCAATGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
 QY 2981 LeuProGluAlaArgLeuAspLeuSerSerTrpPheThrValGlyValAlaGlyIleGly 3000
 DB 9281 TTCGCGAGAGCGGCTCTCTGATTTATCCAGCTGGTTCACCGTCGGGCGGCGGGGCG 9340
 QY 3001 AspIleTyrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu 3020
 DB 9341 GACATTTATCAGAGGCTGTGCGTGGCGGACCGGCTTATGCTCTTTGGCTTACTCTTA 9400
 QY 3021 LeuPheValGlyValGlyIleuPheLeuLeuProAlaArg 3033
 DB 9401 CTTTTTGTAGGGTAGGCTTTTCTTACTCCCGCTCGG 9439
 RESULT 4
 AAQ38218
 ID AAQ38218 standard; cDNA to mRNA; 9589 BP.
 AC AAQ38218;
 XX 25-MAR-2003 (updated)
 DT 01-JUL-1993 (first entry)
 XX NANBH virus strain HC-J6 genome cDNA sequence.
 DE Non A non B hepatitis virus; amplification; HC-J1; HC-J8; plasma; ss.
 XX Non A, non B hepatitis virus strain HC-J6.
 XX Key Location/Qualifiers
 FT CDS 341..9442
 FT /*tag= a
 XX EP532167-A2.
 XX 17-MAR-1993.
 XX 30-JUL-1992; 92EP-0306952.
 PR 09-AUG-1991; 91JP-0287402.
 PR 05-DEC-1991; 91JP-0360441.
 PA (IMMO) IMMUNO JAPAN INC.
 XX Nakamura-T, Okamoto-H;
 WPI, 1993-087166/11.
 DR N-PSDB; AAR33538.
 PT Polynucleotide(s), polypeptide(s) and antibodies of NANBH virus -
 PT useful for detecting NANBH, as a vaccine and for screening blood
 PT samples
 XX Claim 2; Page 27-32; 93pp; English.
 CC RNA was isolated from the plasma of human patients positive for
 CC NANBH virus (strain HC-J6) and was subjected to reverse transcription
 CC to produce cDNA. The resulting cDNA was amplified by PCR, and
 CC nucleic acid sequences determined by analysis of both clones from the
 CC cDNA library and clones obd. by PCR amplification (36 clones in total).
 CC The NANBH HC-J6 genome was found to contain an open reading frame
 CC encoding a polypeptide precursor of 3033 amino acid residues.
 CC See also AAQ38172-221.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 9589 BP: 1968 A: 2820 C: 2635 G: 2166 T: 0 other;

Alignment Scores:

Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.788 Conservatism: 30
 Best Local Similarity: 97.798 Mismatches: 37
 Query Match: 98.168 Indels: 0
 DB: 14 Gaps: 0

US-09-980-559-2 (1-3033) x AAQ38218 (1-9589)

Qy 1 MetSerThrAsnProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 341 ATGAGCACAAATCTAAACCTCAAGAGAAACCAAAAGAAACACCAACCGTGGCCACAA 400
 Qy 21 AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValLysValLysLeuLeuProArgArg 40
 Db 401 GACGTTAAGTTTCCGGGGGGGGCCAGATCGTTGGCGGAGTATACTTGTTCGGGGCAGG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GGGCCCCAGGTGGGTGGTGGCGGCACAGAGGAAGACTTTCGGAGGGGTCCCGAGCCACGTGGA 520
 Qy 61 ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerThrGlyLysProGly 80
 Db 521 AGCGCGCAGCCATCCCTAAGGATCGGCGCTCCACTGGCAAAATCTTCGGGGAAACCAAGGA 580
 Qy 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 Db 581 TACCCCTGGCCCTATACGGGATGAGGGACTCGGCTGGCGAGATGGCTCTGTGCCCC 640
 Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 641 CGAGGTTCGGTCCCTCTTGGGGCCCCATGACCCCCGGCATAGGTTCGCCAACGTGGGT 700
 Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 701 AAGGTTCATGATACCTTAACGTGGCGGTTCGCGACCTCATGCGGGTACATCCCTGCTGTA 760
 Qy 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 761 GGCGCCCGCTCGCGGGCGTGGCCAGAGCTCTCGCGCATGGCTGAGAGTCTGGAGGAC 820
 Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 821 GGCGTTAAATTTGCAACAGGGAGACTTACCCGGTTCCTCTTCTATCTTCTGTGCGCC 880
 Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 881 CTGCTGTCTGTCATCACCACCCCGGTCTCGCTGCCGAAGTGAAGAAACATCAGTACCAGC 940
 Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaAlaVal 220
 Db 941 TACATGGTGACCAACGATGCAACCAATGATAGCATTAACCTGGCCAACTCCAGGCTGCTGTC 1000
 Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1001 CTCAGTCCCCGGGTGGTCCGTGGAGAGAGTGGGGATACATCTCGGTGCTGGATA 1060
 Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1061 CCGGTCTCACCAAGATGTGGCGTGCAGCAGCCCGCGCCCTCACGACAGGGCTTACGGAGC 1120
 Qy 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1121 CACATTCACATGTGTGATGTCCGCCACCGCTCTGCTGCTCTTTTACGTGGGGGACCTC 1180
 Qy 281 CysGlyGlyValMetLeuAlaGlnMetPheIleValSerProGlnHisIleTrpPhe 300
 Db 1181 TCGGTGGGGTATGCTTCGAGCCCAAGATGTTTATGTCTCGCCACAGACCAACATCTGGTTT 1240

Qy 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisArgMetAlaTrp 320
 Db 1241 GTGCAAGACTGCATTTGCTCATCTACCTTGGTACCATCACTGGACACCGCATGGCGTG 1300
 Qy 321 AspMetMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 1301 GACATGATGATGAACCTGTGCTGCCACCGCTACCATGCTTGGCGGTACGCATGGCGTC 1360
 Qy 341 ProGluValIleIlePheIleIleSerGlyAlaHisIleTrpGlyValMetPheGlyLeuAla 360
 Db 1361 CCCGAGGTCAATCATAGACATCATTTGGCGGGGCTCATTTGGGGCGTCAATGCTGGCTTAC 1420
 Qy 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValValIleLeuLeuLeuAlaAlaGly 380
 Db 1421 TACTTCTCTATGAGGAGCGTGGCCAAAGTGGTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 1480
 Qy 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1481 GTGAGCGGCGAACCACCATACCTTGGGGGTTCTACGGCGCATACGCCGACGACCTTACC 1540
 Qy 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1541 GGCATGTTCTCCCTTGGTGGCCAGGAGAAATCCAGCTCATCAACCAATGGCAGTTGG 1600
 Qy 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 Db 1601 CACATCAACCGCACCGCCCTGAATCAATCACTCTTTGCAACACCGGCTCTCTCCGCTCA 1660
 Qy 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1661 CTGTTCTACACCCACAGCTTCACTGTGTCAAGATGTTCGGAACGATGTCCGCTTCGCGC 1720
 Qy 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1721 AGTATCGAGGCGCTTTCGGGTGGGATGGGGCGCTTACCAATATGAGGACCAATGTCCCAAT 1780
 Qy 481 ProGluAspMetArgProTyrCysTrpHisTyrProArgGlnCysGlyValValSer 500
 Db 1781 CCAGGAGTATGAGACCGTATTGCTGGCAGTACCACCAACAGACAGTGTGTGTAGTCTCC 1840
 Qy 501 AlaLysThrValCysGlyProValTyrCysPheThrProSerProValValValGlyThr 520
 Db 1841 GCGAGCTCTGTGTGGGCCCGAGTACTGTTCACCCCGCCAGCCAGTGTGGGTAGC 1900
 Qy 521 ThrAspArgLeuGlyAlaProThrTyrThrTrpGlyGluAsnGluThrAspValPheLeu 540
 Db 1901 ACCGATAGACTTGAGCGCCCTTACACGTGGGGGAGATGAGACAGATGCTCTTCTTA 1960
 Qy 541 LeuAsnSerThrArgProProLeuGlySerTrpPheGlyCysThrTrpMetAsnSerSer 560
 Db 1961 TTGAACAGCACTCGACCCAGCGCAGGGGTTCATGGTTCCGCTGCACTGTGATCACTCCACT 2020
 Qy 561 GlyTyrThrLysThrCysGlyAlaProProCysArgThrArgAlaAspPheAsnAlaSer 580
 Db 2021 GGCTACACCAAGACTTTCGGCGGCACCCCTCTCCCGCATTTACAGCTGACTTCAATGCCAGC 2080
 Qy 581 ThrAspLeuLeuCysProThrAspCysPheArgLysHisArgProAspThrThrTyrLeuLys 600
 Db 2081 ATGAGCTGTGTGGCCCGCCAGGACTGTTTAGAAGCATCTTGATACCACTACATCAAA 2140
 Qy 601 CysGlySerGlyProTrpLeuThrProArgCysLeuIleAspTyrProTyrArgLeuTrp 620
 Db 2141 TGTGGCTCTGGGCGCTGGCTCGCCAGGAGTCCGTGATCGACTACCCCTACAGGCTCTCGG 2200
 Qy 621 HisTyrProCysThrValAsnTyrThrIlePheLysIleArgMetTyrValGlyVal 640
 Db 2201 CATTACCCCTGCACAGTTAACTATACCATCTTCAAAATAAGGATGTATGTGGGGGGGTC 2260
 Qy 641 GluHisArgLeuThrAlaAlaCysAsnPheThrArgGlyAspArgCysAsnLeuGluAsp 660
 Db 2261 GAGCAGAGGCTCAGCGCTGGTGCAATTTCACTCTGTGGGATCTGTGTGCACTTGGAGAC 2320
 Qy 661 ArgAspArgSerGlnLeuSerProLeuLeuHisSerThrThrGluTrpAlaIleLeuPro 680

Db 2321 AGAGACAGAGTCAACCTGCTCTTGTGCTGCACTCCACACGAGTGGGCACTTTACCT 2380
Qy CysSerTyrSerAspLeuProAlaLeuSerThrGlyLeuLeuHisGlnAsnIle 700
Db 2381 TGCACCTACTCGGACCTCGCGGCTTGTGCACTGGTCTTCTCCACCTCCACCAACATC 2440
Qy ValAspValGlnPheMetTyrGlyLeuSerProAlaLeuThrTyrSerIleValArgTyr 720
Db 2441 GTGAGCGTCAATTATGATATGCGCTATACCTGCTCTCAAAATACATGCTCGGATGG 2500
Qy GluTyrValIleLeuLeuPheLeuLeuAlaAspAlaArgValCysAlaCysLeuTyr 740
Db 2501 GAGTGGGTAGTACTCTTATCTGCTCTTAGCGGACGCGAGGTTTGGCGCTGCTTATGG 2560
Qy MetLeuIleLeuLeuGlyGlnAlaGluAlaLeuGluLeuValIleLeuHisAla 760
Db 2561 ATGCTCATCTGTGGGCGAGCGAGCGACCTAGAGAACTTGTCTGTGACGCT 2620
Qy AlaSerAlaAlaSerCysAsnGlyPheLeuTyrPheValIlePhePheValAlaAlaTyr 780
Db 2621 GCGAGCGGAGCTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2680
Qy TyrIleGlyGlyArgValValProLeuAlaThrTyrSerLeuThrGlyLeuTyrSerPhe 800
Db 2681 TACATCAGGGTGGGTAGTCCCTTGGCTACTTATTCCTCACTGGCTATGTGCTCTT 2740
Qy SerLeuLeuLeuAlaLeuProGlnGlnAlaTyrAlaTyrAspAlaSerValHisGly 820
Db 2741 GCGCTACTGCTCTTACGATTCGCCCAACAGGCTTATGCTTATGACGCACTGTATAGT 2800
Qy GlnIleGlyAlaAlaLeuValMetIleThrLeuPheThrLeuThrProGlyTyrIlys 840
Db 2801 CAGTAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2860
Qy ThrLeuLeuSerArgPheLeuTyrTyrLeuLeuLeuLeuLeuGlyAlaAlaMet 860
Db 2861 ACCCTCTCAGCGGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2920
Qy ValGlnGluTyrAlaProMetGlnValArgGlyGlyValArgAspGlyIleIleTyrPala 880
Db 2921 GTCCAGGAGTGGGACACCACTATGCAAGTGGCGGTGGCGGTGGCGGTGGCGGTGG 2980
Qy ValAlaIlePheTyrProGlyValValPheAspIleThrIysTyrTyrLeuLeuAlaValLeu 900
Db 2981 GTGCGCATTTCTGCGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3040
Qy GlyProAlaTyrLeuLeuLeuGlyAlaLeuThrArgValProTyrPheValArgAlaHis 920
Db 3041 GGGCTGCTTATCTCTAAAGGTGCTTTGAAGCGGTGCGGTGCTTCTGTCAGGCGCTCAC 3100
Qy AlaLeuLeuArgMetCysThrMetAlaArgHisLeuAlaGlyGlyValTyrValGlnMet 940
Db 3101 GCTCTACTAAGGATGTGCACCATGTGTAAGGCATCTCGCGGGGGGTAGTGTCCAGATG 3160
Qy AlaLeuLeuAlaLeuGlyArgTyrThrGlyThrIleTyrAspHisLeuThrProMet 960
Db 3161 GTGCTACTAGCCCTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 3220
Qy SerAspTyrAlaAlaSerGlyLeuArgAspLeuAlaValAlaValGluProIleIlePhe 980
Db 3221 TCGGATTTGGGCTGTATATGCTGCGGACTTGGCGGCTTGGCGGCTTGGCGGCTTGG 3280
Qy SerProMetGluTyrIysValIleValTyrGlyAlaGluThrAlaAlaCysGlyAspIle 1000
Db 3281 AGTCCGATGAGAAAAAAGTATCGTCTGGGAGCGGAGACAGCTGCTTGGGGGATATC 3340
Qy LeuHisGlyLeuProValSerAlaArgLeuGlyArgGluValLeuLeuGlyProAlaAsp 1020
Db 3341 TTACACGCACTTCCGCTGCTGGCGGCTTGGCGGAGGCTTCTCTTGGCGGAGCTGAT 3400
Qy GlyTyrThrSerIysGlyTyrSerLeuLeuAlaProIleThrAlaTyrAlaGlnGlnThr 1040

Db 3401 GGCTATACCTCCAGGGGTGGAGTCTTCTGCGCCGCCCATCTACTGCTTATGCCCAGACGACA 3460
Qy ArgGlyLeuLeuGlyThrIleValValSerMetThrGlyArgAspLeuThrGluGlnAla 1060
Db 3461 CCGCGCCCTTTTGGGCAACCATAGTGGTGGAGCATGACGGGGCGCGCACAGACAGAACAGGCC 3520
Qy GlyGluIleGlnValLeuSerThrValThrGlnSerPheLeuGlyThrSerIleSerGly 1080
Db 3521 GGGGAGATTCAAGTCTGTGTCACGGTCACTAGTCTCTTCGCGAACCAACCATCTCGGG 3580
Qy ValLeuTyrThrValTyrHisGlyAlaGlyAsnLeuThrLeuAlaGlySerArgGlyPro 1100
Db 3581 GTCCTATGGAGCTGCTTACCATAGCTGGAGTGGCAACAGACTTAGCGGCTCAGGGGTCGG 3640
Qy ValThrGlnMetTyrSerSerAlaGluGlyAspLeuValGlyTyrProSerProProGly 1120
Db 3641 GTCACACAGATGTACTTCCAGTGTGAGGGGGACTTAGTGGGGTGGCCAGCCCCCGGG 3700
Qy ThrIysSerLeuGluProCysThrCysGlyAlaValAspLeuTyrLeuValThrArgAsn 1140
Db 3701 ACCAAATCTTTGGAGCGGTGCACGTGTGGAGCGGTGCACCTATACCTGGTCCACGCAAC 3760
Qy AlaAspValIleProAlaArgArgArgGlyAspLeuValGlyAlaLeuLeuSerProArg 1160
Db 3761 GCTGATGTCATCCCGCTCGAGAGCGCGGGACAGCGAGGAGGCTACTCTCCCGAGA 3820
Qy ProLeuSerThrLeuIysGlySerSerGlyGlyProValLeuCysProArgGlyHisAla 1180
Db 3821 CCTCTTCCACCTTGAAGGGGTCTCGGGGGGCGCGGTGCTCTGCCCGAGAGGCCACGCT 3880
Qy ValGlyValPheArgAlaAlaValCysSerArgGlyValAlaLysSerIleAspPheIle 1200
Db 3881 GTCGGGTCTTTCGGGCGAGCGGTGCTCTCGGGGGCGGTGGCCAAAGTCCATAGATTTTATC 3940
Qy ProValGlnThrLeuAspIleValThrArgSerProThrPheSerAspAsnSerThrPro 1220
Db 3941 CCGTGTGAGACACTTGCACATCTCACTCGGTCCCCCCTTTAGTGACACAGCACACCA 4000
Qy ProAlaValProGlnThrTyrGlnValGlyTyrLeuHisAlaProThrGlySerGlyIys 1240
Db 4001 CTTGCTGTGCCCCAACTTATCAGGTGGGTACTTACATGCCCCGAGCTGGTAGTGGAAAG 4060
Qy SerThrIysValProValAlaTyrAlaAlaGlnGlyTyrIysValLeuValLeuAsnPro 1260
Db 4061 AGCACCAAAAGTCCCTGTGCGGTATGCGCTCAGGGGTACAAAGTGTCTAGTGTCTTAATCCC 4120
Qy SerValAlaAlaThrLeuGlyPheGlyAlaTyrLeuSerIysAlaHisGlyIleAsnPro 1280
Db 4121 TCGGTGGCTGCCACCTGGGGTGTGGGGGTACTTGTCCAAGGCACATGTCATCAATCCC 4180
Qy AsnIleArgThrGlyValArgThrValThrGlyValAlaProIleThrTyrSerThrTyr 1300
Db 4181 AACATTAGGACTGGGGTCAAGGACTGTGACGACCGGGGGCGCCATCAGTACTCACATAT 4240
Qy GlyIysPheLeuAlaAspGlyGlyCysAlaGlyGlyAlaTyrAspIleIleIleCysAsp 1320
Db 4241 GGCAAATTTCTCGCGCATGGGGGTGGCGAGCGCGCGCTATGACATCATCATATGCGAT 4300
Qy GluCysHisAlaValAspSerThrThrIleLeuGlyIleGlyThrValLeuAspGlnAla 1340
Db 4301 GAATGCCATGCGGTGCACTTACCACCATCTTCCGGCATCGGACAGTCTCTGATCANGCA 4360
Qy GluThrAlaGlyValArgLeuThrValLeuAlaThrAlaThrProProGlySerValThr 1360
Db 4361 GAGACAGCGGGGTCAAGGCTAACTGTACTGGCTACGGCTACGCGCCCCCGGGTCTGTGACA 4420
Qy ThrProHisProAsnIleGluGluValAlaLeuGlyGlnGluGlyGluIleProPheTyr 1380
Db 4421 ACCCCCAACCCCAACATAGAGAGGTGGCCCTCGGGCAGAGGGGTGAGATGCCCTTCTAT 4480
Qy GlyArgAlaIleProLeuSerTyrIleIysGlyGlyArgHisLeuIlePheCysHisSer 1400
Db 4481 GGGAGGGCGATTCCCTCTGTCTATCATCAAGGGAGAGACACTTGTATCTTCTGCACTCA 4540

Db 6701 CCTTCTCCAGAGTCTTTTCTCGGTGGACGGAGTCAGATCCATAGGTTTGCCCCCATTA 6760
 Qy ProlyseProPheAraAspGluValSerPheCyValGlyLeuAenSerPheValVal 2160
 Db 6761 CCGAAGCCGTTTTCGGGATGAGGTCTCGTTCTCGGTTCGGCTTAATTCATTTCTGCTG 6820
 Qy GlySerGlnLeuProCyAspProGluProAspThrAspValLeuMetSerMetLeuThr 2180
 Db 6821 GGGTCTCAGCTTCCTTGGATCTGACCTGACGACGATATGACGTCCATGCTAACA 6880
 Qy AspProSerHisIleThrAlaGluThrAlaAlaArgLeuAlaArgGlySerProPro 2200
 Db 6881 GACCCATCCCATATACCGCGGAGACTCGACGGGGTGTGGACAGGGGTTCACCCCGG 6940
 Qy SerGluAseSerSerAlaSerGlnLeuSerAlaProSerLeuArgAlaThrCysThr 2220
 Db 6941 TCCGAGGCAAGCTCTCAGCGAGCCAGCTATCGGCACCATCGCTCGAGCCACCTGCACC 7000
 Qy ThrHisGlyValAlaTyAspValAspMetValAspAlaAsnLeuPheMetGlyAsp 2240
 Db 7001 ACCCAAGGCAAGCTCTATGATGTGACATGATGTGATGCCAACCTGTTTCATGGGGGGCGAT 7060
 Qy ValThrArgIleGluSerGlySerIleValValValLeuAspSerLeuAspProMetVal 2260
 Db 7061 GTACCCGATAGATCTGAGTCCAAAGTGTGTCTGACCTCTCTCGACCCCAATGTC 7120
 Qy GluLeuArgSerAspLeuGluProSerIleProSerGluTyMetLeuProIleLysArg 2280
 Db 7121 GAAGAAAGGCGACCTTGAGCTCTCGATACCATCGGAATATATGCTCTCCCAAGAGAGA 7180
 Qy PheProAlaLeuProAlaTTPAlaArgProAspTyAsnProLeuValGluSer 2300
 Db 7181 TTCCACAGCCCTTACCGGCTTGGGCAAGGCTGATATCAACCCACCCCTTGTGGATCG 7240
 Qy TrpLysArgProAspTyGlnProAlaThrValAlaGlyCyAlaLeuProProArg 2320
 Db 7241 TGGAGAGCCAGATATACCAACCGCCCACTGTGTGGGGCTGGGCTCTCCCGCCCGCTAAG 7300
 Qy LysThrProThrProProArgArgArgThrValGlyLeuSerGluAspSerIle 2340
 Db 7301 AAAACCCGAGCGCTCCCGCAAGCAGACGCGGACAGTGGGTCTGAGTGAGAGCTCCATA 7360
 Qy GlyAspAlaLeuGlnGlnLeuAlaIleLysSerPheGlyGlnProProSerGlyAsp 2380
 Db 7361 GCAGATGCCCTTACACAGCTGGCCATCAAGTCTTTGGCCAGCCCGCCCGGAGCGGAT 7420
 Qy SerGlyLeuSerThrGlyAlaGlyAlaAlaAspSerGlySerGlnThrProProAspGlu 2380
 Db 7421 TCAGGCTTTTCACGGGGGGGAGCGACGAGCGATTCGGCAGTTCGGACCGCCCGCGATGAG 7480
 Qy LeuAlaLeuSerGluThrGlySerIleSerSerMetProProLeuGluGlyGluLeuGly 2400
 Db 7481 TTGGGCTTTTCGGAGACAGGTTCATCTCTCCATGCCCTCTCGAGGGGGAGCTCGA 7540
 Qy AspProAspLeuGluProGluGlnValGluProGlnProProGlnGlyValAla 2420
 Db 7541 GATCCAGACTTGGAGCTGAGCAGGTAGAGTTCACCTCTCCCGCCAGGGGGGGTGTGA 7600
 Qy AlaProGlySerAspSerGlySerTrpSerThrCysSerGluGluAspSerValVal 2440
 Db 7601 ACCCGGCTCAGGCTCGGGGTCTGGTCTACTTGTCTCGAGGAGGACGACTCCGCTGTG 7660
 Qy CysCysSerMetSerTySerTrpThrGlyAlaLeuIleThrProCysSerProGluGlu 2460
 Db 7661 TGCTGCTCATGTATATCTCTGACCGGGGCTCTAAATACCTCTCTGTAGCCCGGAGAG 7720
 Qy GluLysLeuProIleAsnProLeuSerAsnSerLeuLeuArgTyHisAsnLysValTyR 2480
 Db 7721 GAAAGTTGCCAAATTGGCCCTTGGACCACTCTCTGTTGGATATATCAACAAGGCTGATC 7780
 Qy CysThrThrThrLysSerAlaSerLeuArgAlaLysValThrPheAspArgMetGln 2500

Db 7781 TGTACCAATCAAGAGCGCTCTATTAGGGGCTAAAAAGGTAACTTTTGTATAGATCAAA 7840
 Qy ValLeuAspSerTyTyAspSerValLeuLysAspIleLysLeuAlaAlaSerLysVal 2520
 Db 7841 GGGCTCGAGCGCTCATATGACTCAGTCTTGAAGGACATTAAGCTAGCGGCTCTCCAGGTC 7900
 Qy ThrAlaArgLeuLeuThrMetGluGluAlaCysGlnLeuThrProProHisSerAlaArg 2540
 Db 7901 ACCGCAAGGCTTCTCAGCTTAGAGGAGGCTTCCAGTTAACTCCACCCCACTCTCTGAAGA 7960
 Qy SerLysTyGlyPheGlyAlaLysGluValArgSerLeuSerGlyArgAlaValAsnHis 2560
 Db 7961 TCCAAGTATGGTGTGGGCTAAGGAGGTCCGACGCTTGTCCGGAGAGCGCTTAACCA 8020
 Qy IleLysSerValTrpLysAspLeuLeuGluAspSerGluThrProIleProThrThrIle 2580
 Db 8021 ATCAAGTCCGTGTGGAGGACCTCTCTGGNAGACACAAACACCAATCTCTACAAACATC 8080
 Qy MetAlaLysAsnGluValPheCysValAspProThrLysGlyGlyLysAlaAlaArg 2600
 Db 8081 ATGCCCAAAATGAGGTGTCTCGGTGAGACCCCAACAAAGGGGGGTAAAGAACGAGCTCGC 8140
 Qy LeuIleValTyProAspLeuGlyValArgValCysGluLysMetAlaLeuTyAspIle 2620
 Db 8141 CTATATCGTTTACCTCGACCTCGCGCTCAGGGTCTCGGAGAAAAATGGGCTTTATGATATC 8200
 Qy ThrGlnLysLeuProGlnAlaValMetGlyAlaSerTyGlyPheGlnTySerProAla 2640
 Db 8201 ACACAAAGCTTCTCTCAGGGGTGATGGGGCTCTCTATGATTCAGTACTCTCCCGCT 8260
 Qy GlnArgValGluPheLeuLeuLysAlaTrpAlaGluLysLysAspProMetGlyPheSer 2660
 Db 8261 CAGCGGTGAGGTCTCTTGAAGGATGGGGGAAAGAACCCCTATGGGGTTTTTCG 8320
 Qy TyAspThrArgCysPheAspSerThrValThrGluArgAspIleArgThrGluGluSer 2680
 Db 8321 TATGATACCGGATGCTTTGACTCAACCGCTCACTGAGAGACATCAGSACTGAGGAGTCC 8380
 Qy IleTyArgAlaCysSerLeuProGluGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 Db 8381 ATATATCGGCTTGTCTTCTTCCCGGAGGAGGCCACACTGCCATCACTCACTGACTGAG 8440
 Qy ArgLeuTyValGlyGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyArgArg 2720
 Db 8441 AGACTTTATCTGGAGGGGCCCATGTTCAACAGAGGGCCAGACTCGGGGTACAGGCGT 8500
 Qy CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyValLys 2740
 Db 8501 TGGCGGCCAGCGGGGTGCTTACCACTAGCATGGGGAACACCATCATCTATGTGAAA 8560
 Qy AlaLeuAlaAlaCysLysAlaAlaGlyIleAlaProThrMetLeuValCysGlyAsp 2760
 Db 8561 GCCTTAGCGGCTGTAAAGGCTGACGGGATAATTTGGGCCCAATGCTGTGTATCGGGAT 8620
 Qy AspLeuValValIleSerGluSerGlnGlyThrGluGluAspGluArgAsnLeuArgAla 2780
 Db 8621 GACTTGTGTCTCATCTCAGAGAGCCAGGGGACCCAGAGGAGGAGGAGGAGGAGGAGGAG 8680
 Qy PheThrGluAlaMetThrArgTySerAlaProProGlyAspProProArgProGluTyR 2800
 Db 8681 TTCAAGGAGCTATGACAGGTATTTGCCCCCTCTCTGCTGACCCCCCGGAGGAGGAGGAT 8740
 Qy AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 Db 8741 GACTCGAGCTGATAACATCTTGTCTCTCAATGTGTCTGTGGGCTTGGGCCCGGAGGAG 8800
 Qy ArgArgArgTyTrpLeuThrArgAspProThrThrProIleAlaAlaAlaAlaTrpGlu 2840
 Db 8801 CGCGGAGATACTACTGACGAGAGCCCTTACCACTCCATCGGCCCGGGGTGCTCTGGGAA 8860
 Qy ThrValArgHisSerProValAsnSerTrpLeuGlyAsnIleIleGlnTyAlaProThr 2860
 Db 8861 ACAGTTAGACACTCTCCCTGTCAATTCATGGCTAGGAAACATCATCTAGTACGGCCCAACC 8920

QY 2861 lletroAlaAaMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 DB 8921 ATATGGCTGGATGGTCTGATGACACACTTCTTCATCTCATGAGCCCAAGATACT 8980
 QY 2881 LeuAspGlnAsnLeuAsnPhelMetTyGlyAlaValTySerValSerProLeuAsp 2900
 DB 8981 CTGGACCAAGAACCTCACTTTGATGATGATGAGAGCGGTACTCCGTGAGTCCCTTGGAC 9040
 QY 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTr 2920
 DB 9041 CTCCAGCCCAATATTGAAGGCTTACAGCGGCTTGACGCTTCTCTGACACATACACT 9100
 QY 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 DB 9101 CCCACAGACTGACAGCGGTGCTTCAGCCCTCAGAAACTTGGGGCGCCACCTCCAG 9160
 QY 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyArgAla 2960
 DB 9161 GGTGGAGAGCGCGGACGTCAGTCAAGCGGCTCCCTATCTCCGTGGGGAGAGCG 9220
 QY 2961 AlaValCysGlyArgTyLeuPheAsnTrpAlaValTyThrLysLeuLysLeuThrPro 2980
 DB 9221 GCGTTTGGCGCGATATCTTCAACTGGGCGGTGAAGACCAAGCTCAAACTCACTCCA 9280
 QY 2981 LeuProGluAlaArgLeuAspLysSerTrpPheThrValGlyAlaGlyGly 3000
 DB 9281 TTGCGGAGAGCGCGCTCTCTGGATTATCCAGCTGGTTCACTGTGCGGCGCGCGGCG 9340
 QY 3001 AspileTyHisSerValSerArgAlaArgProArgLeuLeuLeuPheGlyLeuLeu 3020
 DB 9341 GACATTTATCAAGGCTGTGCGTGGCGGACCGCTTATCTCTTGGCTTACTCTCA 9400
 QY 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 DB 9401 CTTTGTGAGGGTAGGCTTCTTCTTCTCCCGCTCGG 9439

RESULT 5
 ID ABK88904 standard; cDNA; 9678 BP.

XX AC ABK88904;
 XX PT 07-OCT-2002 (first entry)
 XX DE Human HCV-related polypeptide encoding cDNA.
 XX KW Hepatitis C virus; human; gene; ss; virucide; gene therapy; HCV;
 XX KW fulminant hepatitis C.
 XX OS Homo sapiens.

Key Location/Qualifiers
 CDS 341..942
 /tag= a
 /product= "Human HCV-related protein"

JP2002171978-A.

18-JUN-2002.

01-DEC-2000; 2000JP-0367365.

01-DEC-2000; 2000JP-0367365.

(TOR-) ZH TOKYO TO RINSHO IGAKU SOGO KENKYUSHO.
 (TORA) TORAY IND INC.

WPI; 2002-569884/61.

P-PSDB; ABG30688.

A gene of a fulminant hepatitis C virus strain and the encoded polypeptide useful in gene therapy to treat hepatitis C.

XX Claim 5; Page 13-25; 36pp; Japanese.
 XX The invention relates to a human polypeptide related to hepatitis C virus (HCV), and the polynucleotide encoding it. The polypeptide can be used for the development of gene therapy on fulminant hepatitis C. This sequence represents cDNA encoding a human HCV-related polypeptide.
 XX Sequence 9678 BP; 1948 A; 2915 C; 2723 G; 2092 T; 0 other;

Alignment Scores:
 Pred. No.: 0 Length: 9678
 Score: 14943.00 Matches: 2765
 Percent Similarity: 95.424 Conservative: 129
 Best Local Similarity: 91.164 Mismatches: 139
 Query Match: 92.704 Indels: 0
 DB: 24 Gaps: 0

US-09-980-559-2 (1-3033) x ABK88904 (1-9678)

QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 DB 341 ATGAGCAAAATCTTAACCTCAAGAAAAACCAAAAGAAACACCAACCTCGCCAGAA 400
 QY 21 AspValLysPheProGlyGlyGlnIleValGlyGlyValTyLeuLeuProArgArg 40
 DB 401 GAGCTTAAGTTCCCGGCGGCGGCGAGATGGTTGGCGGAGTACTTGTTCGCGCGGAG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 DB 461 GCGCCCAAGTGGGTGTGGCGACGACGACGAGAAACCTTCGAGCGGTCCCGACGACG 520
 QY 61 ArgArgGlnProLysProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
 DB 521 AGACGCCAGCCCATCCCAAGATCGGCGCTCACTGCAAGGCGCTGGGAAACCAAGT 580
 QY 81 TyrProTrpProLeuTyArgLysAsnGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 DB 581 CGCCCTGGGCGCTATATGGAATGAGGAGCTGGCTGGCGAGGATGGCTCTGTGCCCC 640
 QY 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 DB 641 CGAGGCTCTGCGCCCTCTCGGGGCCCTCACTGACCCCGGCAATAGGTCCGCGACGTGG 700
 QY 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyTrpIleProValVal 140
 DB 701 AAAGTCATCGACACCTTAACGTGTGGCTTTGCGGACCTCATGGGGTACATCCCGTCTGA 760
 QY 141 GlyAlaProLeuGlyGlyValAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 DB 761 GCGCGCCCGCTTAGTGGCGCGCCAGAGCTGTGCGGCGACGCGGTGAGAGTCTCTGGAGG 820
 QY 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuAla 180
 DB 821 GGGGTAAATATGACACAGGAACTTACCGGTTTCCCGCTTTCTATCTTCTGTGCGCC 880
 QY 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 DB 881 CTGTGTCTGTCATACCGTTCGCGTCTCTGCTGCCAGGTGAGAGATACCAAGTACAGC 940
 QY 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaVal 220
 DB 941 TACATGGTGAACAATGACTCTCAATGACAGCATCACTTGGCAGCTCGAGGCTCGGCTT 1000
 QY 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 DB 1001 CTCCAGTCCCGGCTCGTCCCGTCCGAGAGAGTGGGGATACCTGACGGTGTGGATG 1060
 QY 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 DB 1061 CCAATCTCGCAACATGCGTGTGGCGAGCGCGGCTGCTCAGCGAGGCTCTCGGAGC 1120
 QY 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyTrpValGlyAspLeu 280

```

QY 9361 GGTGTCGACGACCCGCTTATGCTCTTTGGCCCTACTCTCTCTTTTGTAGGGTAGGCT 9420
DB 9361 GGTGTCGACGACCCGCTTATGCTCTTTGGCCCTACTCTCTCTTTTGTAGGGTAGGCT 9420
QY 9421 TTTCTACTCTCCGCTCGGTAGAGCGGACACATTAGCTACACTCCATAGCTTAACCTGCC 9480
DB 9421 TTTCTACTCTCCGCTCGGTAGAGCGGACACATTAGCTACACTCCATAGCTTAACCTGCC 9480
QY 9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
DB 9481 CTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9540
QY 9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9600
DB 9541 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9600
QY 9601 TACTTTCTTCTGTGCTCCATCTTTAGCCCTAGTCAAGCTAGCTGTGAAGGTCCGT 9660
DB 9601 TACTTTCTTCTGTGCTCCATCTTTAGCCCTAGTCAAGCTAGCTGTGAAGGTCCGT 9660
QY 9661 GAGCGCATGACTGACAGAGTCCGTAACCTGCTCTCTGACATCATGT 9711
DB 9661 GAGCGCATGACTGACAGAGTCCGTAACCTGCTCTCTGACATCATGT 9711

RESULT 3
AX057395
LOCUS AX057395 9711 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 4 from Patent WO0075337.
ACCESSION AX057395
VERSION AX057395.1 GI:12310134
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
VIRUSES: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
1
REFERENCE
AUTHORS Bukh, J., Yanagi, M., Emerson, S.U. and Purcell, R.H.
TITLE Infectious cDNA clone of gb virus b and uses thereof
JOURNAL Patent: WO 0075337-A 4 14-DEC-2000;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES
Location/Qualifiers
1..9711
/organism="Hepatitis C virus"
/mol_type="genomic DNA"
/db_xref="taxon:11103"
BASE COUNT 1987 a 2853 c 2668 g 2203 t
ORIGIN
Query Match 100.0%; Score 9711; DB 6; Length 9711;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCCGCCCTAATAGGGGACACTCCGCAATGAATCACTCCCTGTGAGGAACTACTGT 60
DB 1 ACCCGCCCTAATAGGGGACACTCCGCAATGAATCACTCCCTGTGAGGAACTACTGT 60
QY 61 CTTACAGCAGAAAGCGCTAGCCATGCGTATGATGATGCTGATGATGCTGATGATGCTG 120
DB 61 CTTACAGCAGAAAGCGCTAGCCATGCGTATGATGATGCTGATGATGCTGATGATGCTG 120
QY 121 CCCCTCCCGGAGAGCCATAGTGTCTGCGAACCCTGAGTACACCGGAATTCGCGG 180
DB 121 CCCCTCCCGGAGAGCCATAGTGTCTGCGAACCCTGAGTACACCGGAATTCGCGG 180
QY 181 AAGACTGGTCTCTTTCTTGATTAACCACTATGCGCCGCCATTTGGGCGTGCCTCCG 240
DB 181 AAGACTGGTCTCTTTCTTGATTAACCACTATGCGCCGCCATTTGGGCGTGCCTCCG 240
QY 241 CAAGACTGCTAGCGAGTAGCTGTGGTTGCGAAGGCTTGTGGTACTCTCTGATAGG 300
DB 241 CAAGACTGCTAGCGAGTAGCTGTGGTTGCGAAGGCTTGTGGTACTCTCTGATAGG 300

```

```

QY 301 TGCTTGCGAGTGCCTCGGAGGTCTCGTAGACCGTGCACCATGAGCACAATCTTAAACC 360
DB 301 TGCTTGCGAGTGCCTCGGAGGTCTCGTAGACCGTGCACCATGAGCACAATCTTAAACC 360
QY 361 TCAAAGAAAACCAAAAGAAAACCAACCGCTCGCCCAACAGAGCTTAAGTTTCGGGCGG 420
DB 361 TCAAAGAAAACCAAAAGAAAACCAACCGCTCGCCCAACAGAGCTTAAGTTTCGGGCGG 420
QY 421 CGGCCAGATCGTTGGCGAGTATATCTTGTGCGCGCAGGGGCCCAAGTTGGTGTGG 480
DB 421 CGGCCAGATCGTTGGCGAGTATATCTTGTGCGCGCAGGGGCCCAAGTTGGTGTGG 480
QY 481 CGGCACAAGGAAGACTTCGGAGCGGTCCAGCCACAGTGAAGCGCCAGCCATCCCTAA 540
DB 481 CGGCACAAGGAAGACTTCGGAGCGGTCCAGCCACAGTGAAGCGCCAGCCATCCCTAA 540
QY 541 AGATCGGCGCTCCACTTGGGGAATCTTGGGGAACACAGGATACCCCTGGGCCCTTAT 600
DB 541 AGATCGGCGCTCCACTTGGGGAATCTTGGGGAACACAGGATACCCCTGGGCCCTTAT 600
QY 601 GAATGAGGAGCTCGGCTGGGAGATGCTCTCTGTCGCCCGAGGTTCCGTCCTCTTG 660
DB 601 GAATGAGGAGCTCGGCTGGGAGATGCTCTCTGTCGCCCGAGGTTCCGTCCTCTTG 660
QY 661 GGGCCCAAATGAACCCCGCATAGGTCCGCAACAGTGGTAAGGTCAATCGATACCCCTAAC 720
DB 661 GGGCCCAAATGAACCCCGCATAGGTCCGCAACAGTGGTAAGGTCAATCGATACCCCTAAC 720
QY 721 GTGCGGCTTTGGCGACCTCATCGGGTACATCCCTGCTGGGCGCCCGCTCGGCGGCT 780
DB 721 GTGCGGCTTTGGCGACCTCATCGGGTACATCCCTGCTGGGCGCCCGCTCGGCGGCT 780
QY 781 GCGCAGACTCTCGGCGCATGCGGTGAGAGTCTGCGAGACGGGTTAAATTTGCAACAG 840
DB 781 GCGCAGACTCTCGGCGCATGCGGTGAGAGTCTGCGAGACGGGTTAAATTTGCAACAG 840
QY 841 GAACTTACCCGCTTCTCTTTTCTATCTTCTGCTGCGCCCTCTCTCTCTGATACCCAC 900
DB 841 GAACTTACCCGCTTCTCTTTTCTATCTTCTGCTGCGCCCTCTCTCTCTGATACCCAC 900
QY 901 CCGGCTCTCGGCTCGGGAAGTGAAGAACATCAGTACCGGCTACATGCTGATCAACGACTG 960
DB 901 CCGGCTCTCGGCTCGGGAAGTGAAGAACATCAGTACCGGCTACATGCTGATCAACGACTG 960
QY 961 CACCAATGACAGCATTTACTTGGCAGCTCCAGGCTGCTGCTCCAGCTCGCGGCTGCT 1020
DB 961 CACCAATGACAGCATTTACTTGGCAGCTCCAGGCTGCTGCTCCAGCTCGCGGCTGCT 1020
QY 1021 CCGGTGCGAAGTGGGGAATGATCTCAGTGTCTGATACCGTCTCACCGAATGTGGC 1080
DB 1021 CCGGTGCGAAGTGGGGAATGATCTCAGTGTCTGATACCGTCTCACCGAATGTGGC 1080
QY 1081 CGTGACGAGCGCCGCGCCCTCAGCGAGGCTTTCGGAACGCAATCGATCGATGCTGAT 1140
DB 1081 CGTGACGAGCGCCGCGCCCTCAGCGAGGCTTTCGGAACGCAATCGATCGATGCTGAT 1140
QY 1141 GTCCGCGAGCTCTGCTCTGCTCTTACGTGGGGAGCTCTGCGGTGGGCTGATCTGCT 1200
DB 1141 GTCCGCGAGCTCTGCTCTGCTCTTACGTGGGGAGCTCTGCGGTGGGCTGATCTGCT 1200
QY 1201 AGCCCAATGTTTCAATGTTCTGCGCGAGCACTCTGTTTGTCCAAAGATGCAATTTGCTC 1260
DB 1201 AGCCCAATGTTTCAATGTTCTGCGCGAGCACTCTGTTTGTCCAAAGATGCAATTTGCTC 1260
QY 1261 CATCTACCCTGGTACCACTACTGGACACCGCATGGGATGGGATGATGATGAACTGGTCT 1320
DB 1261 CATCTACCCTGGTACCACTACTGGACACCGCATGGGATGGGATGATGATGAACTGGTCT 1320
QY 1321 GCGCAGCGCTACCATGATCTTGGGCTAGCGATGCGTGTCCCGAGGCTCATTTATAGAT 1380
DB 1321 GCGCAGCGCTACCATGATCTTGGGCTAGCGATGCGTGTCCCGAGGCTCATTTATAGAT 1380

```


ORGANISM Hepatitis C virus
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 1 (bases 1 to 9589)
 Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A., Miyakawa, Y. and Mayumi, M.
 Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions
 J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
 92044440
 MEDLINE
 PUBMED
 1658196
 REFERENCE
 2 (sites)
 Han, J. H. and Houghton, M.
 Group specific sequences and conserved secondary structures at the 3' end of HCV genome and its implication for viral replication
 Nucleic Acids Res. 20 (13), 3520 (1992)
 92335016
 MEDLINE
 PUBMED
 1321416
 REFERENCE
 3 (sites)
 Hotta, H., Doi, H., Hayashi, T., Purwanta, M., Soemarto, W., Mitokami, M., Ooba, K. and Homma, M.
 Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia
 Arch. Virol. 136 (1-2), 53-62 (1994)
 94270990
 MEDLINE
 PUBMED
 7545932
 COMMENT
 These data kindly submitted in computer readable form by: Hiroaki Okamoto
 Immunology Division
 Jichi Medical School
 Kawachi-gun
 Tochigi-Ken 329-04
 Japan
 Phone: 0285-44-2111 x3334
 Fax: 0285-44-1557.
 Location/Qualifiers
 1. 9589
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="HC-J6"
 /db_xref="taxon:11103"
 341. 9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="BA00792.1"
 /db_xref="GI:211651"
 /translation="MSTNPKPQKTKRNNRRPDQVFPGGQIVGGVYLLPRGPRLLGVRAKTBRSOPRRQPIPKRRSTGKSGKPGYPWPLXGNEGLWAGWLLSPRGSRPSGPNDRHSRNVKGVDTLTCGPADLMGYIPVVGAPLGGVAAALAHGVLEDAVNFATGNLPGCSFIFILLALSCITTPVSAEYKNIISTGYMTNDCTNDSITWLOQAVLHVPCCVPCVKGNTRSRCPVSPNVAQVQPGALTOGLRTHDMVMSATLCSALYVGLCGVMIAAQMEFVSPQHMFVQDCNCSIYPGTITGHRMAMDMWMSPTAMILAVAMRPVPIIDIIIGAHGVNFGVLAVPSMOGAWKVVVILLAAQVDAQTHVGSTAHNARTLGMFSLGAROKIOLINTNGSHINRTALNCNDSLHGTPLASLPTHSFNSSGCPERSACRSLEAFRVGMALQYEDVTHPEDMRPYCHYPRQGVVSASVCGPVYCTPSPPVVTGTRDLGAPTYTNGENRTHVFLNSTRPPGCSFPGCTNNSTGYTKTGAPPCIRADFNASMDLLCPTDCPKRPDITYIKCGSGFPLRCLIDIPYKLMHYPCYNYTIFKRLMTVGSRHRTAACNFTGRNCDNRDSQSPLHSTTEWAILPCTYSLPALTGLLHLHONIIVDQFMIGSPALTKYIWRNEWVLLFLLADARVCACLWMLLIGQAALALKLVHLHSAASACNGFLYFVFFVAAMYIKRNVPLATYSLGLWSPGLLALPOQAYDASVHGOIGALLVLTPTLTGKYLTLRSFLMLCYLLTLAEAMQEWAPMVGQGRDGIIMAVAI PCFQVVDITKWLAVLAGPAILLKGALTRVPYVFAHALRMCMTYVHLGGRVYQVLLALGRWTGTIYDHLTPMSMAANGRLDLAVAFPIIFSPMKYIVMGAETAAACGDI LHLGPVSLARGLREVLGADGYTSKGSLLAPITAYAQTRGLLGTIYVSMTRGDKTEQAGEIQVLSVTQSPFLGTTISGLVLTVYHAGNKTLAGSRPVTOMYSSAEGDLVGMFSPGPKTSLEPCTCGAVDLXVTRNADVIPA RRRGDKRALLSPRLSTLKGSSGVPVLCPRHAGVGFRAAVCSRGVAKSIDPFPVETLDIVTRSPVNDSTPPAVPOTYQVGLHAPGTGSKTKVPVAVAGVYKVLINLNSVLAFLGPAVLLKANGINPINTGYVTVTGAPLITYSTYKELADGCGAGYDIIICDECHAVDSTLLGLTVDLQDAETAGVLTATATATPGSVTTTFPHNIEEVALGQGEIIPFYGRALPLSYLKGSHLIFCHSKKKCDLAAALNGMAVATIRGLDVSIVTPQGDV VVATDALMTFTGDFDSVIDCNVAVTQVVDFSLDPTFTITITQTVFQDAVSRQRGR

TCGRLGIVRYVSTGCRASGHPDPSVVICEDDAGANWELTPTAETTVRLRAVETPTPL
 PVCKRLEFWAEVFTGLTHIDHFLSQTKSGENFAYLTATQVACARAKAPPPSDV
 MCKLRLKPTLVGFTPLLYLRGLSVTQKSTVETLTHPVTKYIATQMOADEVMTVWLAG
 GLAAAVAKCQGLATQCVCIIGRLHYNQRAVAPKVEVLEAFDEMECECAALISEGO
 RIAEMUKSICQLAQKQASQKQODIOPAVQASVPKQFQWAKHMFNFSIGIQTLAGIST
 LPNPANVAMFAFSAALTSPLSTSTTLLMLAGWLASQIAPAGATGFPVSVGLVGA
 VSGILGKVLVDILAGYGAGISGALVAFKMSGKPSMEDVNNLLPGLISGALVGVV
 ICALRLRHVIGPEGAVOMVRLIAFASRGHNVAPHYTHVETSDASORVTLQGLSTIT
 SLRLRLHNIITDCPIPCSGSLRWLDVMDVCTILITDKNWLTSKLPKMPGLPFTISQ
 KGVKGVAGTGMITTRCPGANISGNVLSGMRITGPKTKMNTWQTPINCYTSGQC
 VPKAPNPKIALMRVAASEYAEVTHGSHYITGLTDLNKKVPCOLPSPSPSWUDVQ
 QIHRFAPIKPPEFDEVSFCVGLANSFVVSQSLCDPEPDDVLTSLTSPSHITETA
 ARELAGSPSPSEASSASQLSAPSLRATCTHKGADVDMDVNDLVNLPVSVESKRPDY
 KVVLDLDPVVEERSDLEPSIPEYMLPKRPPALPAPWAPDINPPVIVESWKPDPY
 QVATVAGCLAPPPKATLPPTPPRRRTVGLSESSIAKLOOLAISFGQPPPGDGLS
 TGAADSGRTPPDDELALSETGISMSPEEGEPDGDLEPEQVLPQPPGVVTP
 GSGSGSWTCSSEEDSVVCCSMYSWTGALITPCSPSEKLPINPLNSLLRHNKYV
 CITTSASLRKKTVDPMQALDAHYDVLKDIKLAASKVTARLLTLEAQCILTPHVS
 ARSKYFGAKBVRSLGSAVNHKSVNKLLLEDITQPIPTTMAKNEVECVDTKGGK
 KAARLIVPDGLVRCMALYDITQKLPQAVMGASYGFOYSPAQVFEFLKMAWEKK
 DPMFSDYTRCFDSTVTTERDIRTEESYIRACSLPBEAHTAIHSLVYCGPMSNSK
 GQTCYRRCRASGLVITSMGNTITCYVKALACAAAGIITAPMLVCGDLVVISSQG
 TDEDRNLRAFTTEAMTRYSPAPGDPPEPYDLELITSCSSNVSVAGPQGRERYLTR
 DPTPIAAWETVHSPVSNWLCNIIIOYAPTIWARMVLMTHFFSILMAQDRLDONLA
 FEMYGAVYSVPLDLPALIELRLGLDAFSLHTYTPHELTRVASLRKLKGLAPLRWKS
 RARAVASLISRGGRACVCGVLPNVAETKELPLPPEARLLDLSNFTVAGGGDI
 YHSVSRAPRRLILLGLLLLVGVGLFLLPAR*
 BASE COUNT 1970 a 2820 c 2633 g 2166 t
 ORIGIN
 Query Match 92.4%; Score 9973.6; DB 14; Length 9589;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 9204; Conservative 0; Mismatches 384; Indels 0; Gaps 0;
 QY 1 ACCCGGCCCTAATAGCGCGACACTCCGCCATGAATCACTCCCTCTGAGGAACACTACTGT 60
 DB 1 ACCCGGCCCTAATAGCGCGACACTCCGCCATGAATCACTCCCTCTGAGGAACACTACTGT 60
 QY 61 CTTACGAGAGAGAGGCTTACGCGCATGCGTGTGATGAGTGTGTACAGCTCCAGGCC 120
 DB 61 CTTACGAGAGAGAGGCTTACGCGCATGCGTGTGATGAGTGTGTACAGCTCCAGGCC 120
 QY 121 CCGCTTCCCGGAGAGCCTATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGG 180
 DB 121 CCGCTTCCCGGAGAGCCTATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGGG 180
 QY 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCGCCATTTGGCGCTGCCCCCG 240
 DB 181 AAGACTGGGTCTTTCTTGGATAAACCCACTCTATGCCCGCTGCCCCCG 240
 QY 241 CAAGACTGCTACCGCGAGTAGCGTTGGGTTCGGAAGCGCTTGTGCTACTGCTGATAGG 300
 DB 241 CAAGACTGCTACCGCGAGTAGCGTTGGGTTCGGAAGCGCTTGTGCTACTGCTGATAGG 300
 QY 301 TCGTTCCGAGTCCCGCGGAGTCTCTGATAGCGTGTGACCATGAGCAAAATCTCTAAACC 360
 DB 301 TCGTTCCGAGTCCCGCGGAGTCTCTGATAGCGTGTGACCATGAGCAAAATCTCTAAACC 360
 QY 361 TCAAGAGAAAAACCAAAAGAAACCAACCGTCCGCCCAAGACGTTTAACTTTCCGCGGG 420
 DB 361 TCAAGAGAAAAACCAAAAGAAACCAACCGTCCGCCCAAGACGTTTAACTTTCCGCGGG 420
 QY 421 CGGCGCAGATCGTTGGCGGAGTATCTTGTGTCGCGCGACGGGGCCAGGTGCGGTGCG 480
 DB 421 CGGCGCAGATCGTTGGCGGAGTATCTTGTGTCGCGCGACGGGGCCAGGTGCGGTGCG 480
 QY 481 CGGCGCAGAGAGACTTTCGAGCGGTCTCCAGCCACTGGAAGCGCGCAGCCCATCCCTAA 540
 DB 481 CGGCGCAGAGAGACTTTCGAGCGGTCTCCAGCCACTGGAAGCGCGCAGCCCATCCCTAA 540
 QY 541 AGATCGCGCTCTCACTCGCAAAATCTCTGGGAAAAACCAAGATACCCCTCGCCCTATACGG 600
 DB 541 AGATCGCGCTCTCACTCGCAAAATCTCTGGGAAAAACCAAGATACCCCTCGCCCTATACGG 600

8581 TGCAGGAGTATCGCCGACCAATGCTGTATGCGGAGTGAATCTGGTTGTATCTCAGA 8640
 8641 AACCCAGGAGGACCGAGGAGGACGAGACCTGTAGAGCCTTCAACGAGGCTATGACAG 8700
 8641 AACCCAGGAGGACCGAGGAGGACGAGACCTGTAGAGCCTTCAACGAGGCTATGACAG 8700
 8701 GTATTCTGCCCCCTCTCTGTGTGACCCGAGACCGGAGTATGATCTCGAGCTGATACATC 8760
 8701 GTATTCTGCCCCCTCTCTGTGTGACCCGAGACCGGAGTATGATCTCGAGCTGATACATC 8760
 8761 TTGCTCTCAAAATGTCGTGTGGGCTGTGGGCGGACAGGCGGCGGAGATACCTGAC 8820
 8761 TTGCTCTCAAAATGTCGTGTGGGCTGTGGGCGGACAGGCGGCGGAGATACCTGAC 8820
 8821 CAGAGACCTTACACTTCOAATGCGCGGCTGTGGGAAACAGTTAGACACTCCCTCTGT 8880
 8821 CAGAGACCTTACACTTCOAATGCGCGGCTGTGGGAAACAGTTAGACACTCCCTCTGT 8880
 8881 CAATTCAATGCTGGGAAACATCATCCAGTAGCGCCCGGACCATATGCGCTGCGATGCTCT 8940
 8881 CAATTCAATGCTGGGAAACATCATCCAGTAGCGCCCGGACCATATGCGCTGCGATGCTCT 8940
 8941 GATGACACACTTCTCTCATCTCATGCTCATGAGACGCTGAGACCTCACTT 9000
 8941 GATGACACACTTCTCTCATCTCATGCTCATGAGACGCTGAGACCTCACTT 9000
 9001 TGAGATGACGAGGCGGTGACTCCGTTGAGTCCCTTGGACCTCCAGCTATATGAAAG 9060
 9001 TGAGATGACGAGGCGGTGACTCCGTTGAGTCCCTTGGACCTCCAGCTATATGAAAG 9060
 9061 GTTACATGCGCTGTGACGCTTTTCTCTGACACATACACTCCCGACCACTGACAGCGGT 9120
 9061 GTTACATGCGCTGTGACGCTTTTCTCTGACACATACACTCCCGACCACTGACAGCGGT 9120
 9121 GCTTCAGCCCTCAGAAACTTGGGCGCCACCTCTCAGAGCTGAGAGCGGCGGACG 9180
 9121 GCTTCAGCCCTCAGAAACTTGGGCGCCACCTCTCAGAGCTGAGAGCGGCGGACG 9180
 9181 TGCACTCAGGCGGCTCTCATCTCCGTGGGCGGAGGCGGCTTGGCTGATATCT 9240
 9181 TGCACTCAGGCGGCTCTCATCTCCGTGGGCGGAGGCGGCTTGGCTGATATCT 9240
 9241 CTTCAATTGGGCGGTGAGACCAAGCTCAAACTCACTTCCATGCGGAGCGGCGCTCT 9300
 9241 CTTCAATTGGGCGGTGAGACCAAGCTCAAACTCACTTCCATGCGGAGCGGCGCTCT 9300
 9301 GGATTTATCCAGCTGCTTCACTCCGTGGGCGGCGGCGGCGGACATTTATCAGCGGTC 9360
 9301 GGATTTATCCAGCTGCTTCACTCCGTGGGCGGCGGCGGCGGACATTTATCAGCGGTC 9360
 9361 GCGTGCCGACCCGCTTATGCTCTTTGGCTTACTCTTCTTTTGTAGGGGTAGGCT 9420
 9361 GCGTGCCGACCCGCTTATGCTCTTTGGCTTACTCTTCTTTTGTAGGGGTAGGCT 9420
 9421 TTTCCTACTCCCGCTCGGTAGAGCGGCAACAATTAGCTTACATCTCCTCACTAGCTGTC 9480
 9421 TTTCCTACTCCCGCTCGGTAGAGCGGCAACAATTAGCTTACATCTCCTCACTAGCTGTC 9480
 9481 CTTT 9540
 9481 CTTT 9540
 9541 TT 9600
 9541 TT 9600
 9601 TACTTTCTTCTGGTGGCTCCATCTTACGCTTACGCTAGCTGTGAAAGGTCCT 9660
 9601 TACTTTCTTCTGGTGGCTCCATCTTACGCTTACGCTAGCTGTGAAAGGTCCT 9660
 9661 GAGCGCATGCTGACAGAGTGGCGTAACTGCTCTCTCTCAGATCATGT 9711

9661 GAGCGCATGCTGACAGAGTGGCGTAACTGCTCTCTCTCAGATCATGT 9711

RESULT 3
 ID AAF23486
 XX AAF23486 standard; DNA, 9711 BP.
 XX AAF23486;
 DT 21-MAR-2001 (first entry)
 XX Infectious Hepatitis C virus 2a genotype.
 XX GBV-B; hepatitis C virus; HCV; vaccine; ds.
 XX Hepatitis C virus.
 OS OS
 PN WO200075337-A1.
 XX 14-DEC-2000.
 PF 02-JUN-2000; 2000WO-US15293.
 PR 04-JUN-1999; 99US-0137694.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Bukh J, Yanagi M, Emerson SU, Purcell RH;
 WPI; 2001-091214/10.
 XX New infectious nucleic acids of the GB virus-B clone, useful for indirectly studying the molecular properties of hepatitis C virus (HCV) and in developing vaccines and therapeutics for HCV -
 XX Disclosure; Page 78-82; 96pp; English.
 XX The present invention relates to GB virus-B. The nucleic acid molecules of the invention are useful for indirectly studying the molecular properties of hepatitis C virus (HCV). The infectious nucleic acid sequence of the GB virus-B clone and the HCV/GBV-B chimera may be used in the development of vaccines and therapeutics for HCV.
 XX Sequence 9711 BP; 1989 A; 2851 C; 2668 G; 2203 T; 0 other;

Query Match
 Best Local Similarity 100.0%; Score 9707.8; DB 22; Length 9711;
 Matches 9709; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACCGCGCCCTAATAGCGGCGACACTCCGCAATGAATCACTCCCTGTGTGAGGAACTACTGT 60
 DB 1 ACCGCGCCCTAATAGCGGCGACACTCCGCAATGAATCACTCCCTGTGTGAGGAACTACTGT 60
 QY 61 CTTTCAGCGAGAAAGGCTCTAGCCATGGGCTTAGTATGAGTGTCTACAGCTCCAGGCC 120
 DB 61 CTTTCAGCGAGAAAGGCTCTAGCCATGGGCTTAGTATGAGTGTCTACAGGCCCTCCAGGCC 120
 QY 121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGGAACCGGTGAGTACACCGAATGCGGG 180
 DB 121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGGAACCGGTGAGTACACCGAATGCGGG 180
 QY 181 AGAGCTGGGTCCTTTCTTGGATAAACCCACTCTATATGCGCGGCCATTTGGGCGTGC 240
 DB 181 AGAGCTGGGTCCTTTCTTGGATAAACCCACTCTATATGCGCGGCCATTTGGGCGTGC 240
 QY 241 CAAGACTGTAGCTAGCGTAGGCTTGGGTTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
 DB 241 CAAGACTGTAGCTAGCGTAGGCTTGGGTTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
 QY 301 TGTCTGCGAGTGCCTCCGCGGAGTCTCTGAGCCGTGACCATGACGACCAATCTTAACC 360
 DB 301 TGTCTGCGAGTGCCTCCGCGGAGTCTCTGAGCCGTGACCATGACGACCAATCTTAACC 360

Db 2521 CTTGCTTTAGCGGAGCGGCTTGGGCTGCTTATGAGTCTCATCTTGTGGGCA 2580
 Qy 2581 GGCAGGAGCAGCACTAGAGAGCTGCTCATCTTGCAGCTGCGAGCGAGCTAGCTGCAA 2640
 Db 2581 GGCAGGAGCAGCACTAGAGAGCTGCTCATCTTGCAGCTGCGAGCGAGCTAGCTGCAA 2640
 Qy 2641 TGGCTCTCTATATTTGTCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 Db 2641 TGGCTCTCTATATTTGTCATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2700
 Qy 2701 CCGCTTGTAGTACTTATCT 2760
 Db 2701 CCGCTTGTAGTACTTATCT 2760
 Qy 2761 GCGCCAAAGGCTTATGCTTATGAGCGATCTGTCATGCTGCTGCTGCTGCTGCTGCTGCT 2820
 Db 2761 GCGCCAAAGGCTTATGCTTATGAGCGATCTGTCATGCTGCTGCTGCTGCTGCTGCTGCT 2820
 Qy 2821 GGTATGATCACTCTCTTACTCTCACCCCGGCTATAGACCTTCTCAGCCGCTTTT 2880
 Db 2821 GGTATGATCACTCTCTTACTCTCACCCCGGCTATAGACCTTCTCAGCCGCTTTT 2880
 Qy 2881 GTGGTGTGCTATCTTCTGACCTCTGGGGAAGCTATGGTCCAGGAGTGGGCAACCAC 2940
 Db 2881 GTGGTGTGCTATCTTCTGACCTCTGGGGAAGCTATGGTCCAGGAGTGGGCAACCAC 2940
 Qy 2941 TATGAGGTGGCGGCTGATGTCATCATATGGCGCTGCGCATATTTCTACCCAGG 3000
 Db 2941 TATGAGGTGGCGGCTGATGTCATCATATGGCGCTGCGCATATTTCTACCCAGG 3000
 Qy 3001 TGTGGTGTGACATACCAAGTGGCTTGGGGGCTTGGGGCTTGGGGCTTGGGGCTTGGGG 3060
 Db 3001 TGTGGTGTGACATACCAAGTGGCTTGGGGGCTTGGGGCTTGGGGCTTGGGGCTTGGGG 3060
 Qy 3061 AGGTGCTTGTGACCGCGTGGCTGCTTGGCTGAGGGCTCACTCTCTCTCTCTCTCTCTCT 3120
 Db 3061 AGGTGCTTGTGACCGCGTGGCTGCTTGGCTGAGGGCTCACTCTCTCTCTCTCTCTCTCT 3120
 Qy 3121 CATGCAAGGCACTCTCGCGGGGCGAGGTACGCTGAGATGGGCTTACGCTCTCTCTCTCT 3180
 Db 3121 CATGCAAGGCACTCTCGCGGGGCGAGGTACGCTGAGATGGGCTTACGCTCTCTCTCTCT 3180
 Qy 3181 GTGGAGTGGCACTTATGATGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTAC 3240
 Db 3181 GTGGAGTGGCACTTATGATGACCACTTACCTTACCTTACCTTACCTTACCTTACCTTAC 3240
 Qy 3241 CCGTGGGAGCTTGGCGGCTGCGCTTGAAGCTTATCATCTTCTGAGTCCGATGGAGAAAGT 3300
 Db 3241 CCGTGGGAGCTTGGCGGCTGCGCTTGAAGCTTATCATCTTCTGAGTCCGATGGAGAAAGT 3300
 Qy 3301 CATTTGCTGGGAGCGGAGACAGCTGCTTGGGGGACATTTTACACGGGACTTCCCGTGT 3360
 Db 3301 CATTTGCTGGGAGCGGAGACAGCTGCTTGGGGGACATTTTACACGGGACTTCCCGTGT 3360
 Qy 3361 CGCCGAGCTTGGTGGGAGGCTCTCTTGGCCCGGCTATGAGTCTATACCTCCAGGGGTG 3420
 Db 3361 CGCCGAGCTTGGTGGGAGGCTCTCTTGGCCCGGCTATGAGTCTATACCTCCAGGGGTG 3420
 Qy 3421 GAGTCTTCTGCGCCCACTCTGCTTACCGCAGAGACAGCTGCTGCTTGGGCACTTACCA 3480
 Db 3421 GAGTCTTCTGCGCCCACTCTGCTTACCGCAGAGACAGCTGCTGCTTGGGCACTTACCA 3480
 Qy 3481 AGTGTGAGCATGACGGGGGCGGCAAGAGACAGAGTGGGGGAAATTCAGGCTCTGTC 3540
 Db 3481 AGTGTGAGCATGACGGGGGCGGCAAGAGACAGAGTGGGGGAAATTCAGGCTCTGTC 3540
 Qy 3541 CACAGTCACTCAGTCT 3600
 Db 3541 CACAGTCACTCAGTCT 3600
 Qy 3601 TGGAGCTGGCAAGAGCTCTGCGCGGCTCAAGGGTGGGCTGAGGAGTATCTCCAG 3660
 Db 3601 TGGAGCTGGCAAGAGCTCTGCGCGGCTCAAGGGTGGGCTGAGGAGTATCTCCAG 3660

Qy 3661 TGCTGAGGGGAGCTTATGAGTGGGCTGCGCGGCTGCGGACTAAATCTTTTGGAGCGGTG 3720
 Db 3661 TGCTGAGGGGAGCTTATGAGTGGGCTGCGCGGCTGCGGACTAAATCTTTTGGAGCGGTG 3720
 Qy 3721 CAGTGTGGAGCGGCTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
 Db 3721 CAGTGTGGAGCGGCTGCTGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3780
 Qy 3781 AAGACGCGGGGACAAACGCGGGAGCGCTACTCTCTCCCGGAGACTCTTTTCCACCTTGAAGG 3840
 Db 3781 AAGACGCGGGGACAAACGCGGGAGCGCTACTCTCTCCCGGAGACTCTTTTCCACCTTGAAGG 3840
 Qy 3841 GTCCTCAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
 Db 3841 GTCCTCAGGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3900
 Qy 3901 TGTGTGCTCTCGGGGGTGGCTTAACTCATAGATTTATCTCCCGTGGAGACACTCGACAT 3960
 Db 3901 TGTGTGCTCTCGGGGGTGGCTTAACTCATAGATTTATCTCCCGTGGAGACACTCGACAT 3960
 Qy 3961 CGTCAGCGGCTCCCGCACCTTTAGTGACACAGCACACCACTGCTGCTGCTGCTGCTGCTGCTGCT 4020
 Db 3961 CGTCAGCGGCTCCCGCACCTTTAGTGACACAGCACACCACTGCTGCTGCTGCTGCTGCTGCTGCT 4020
 Qy 4021 TCAGTGGGCTTACTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
 Db 4021 TCAGTGGGCTTACTTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4080
 Qy 4081 ATATGCTGCTCAGGGGTATTAAGTCTGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 4140
 Db 4081 ATATGCTGCTCAGGGGTATTAAGTCTGCTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 4140
 Qy 4141 GTTGGGGGCTTACTTGTCTTAAAGGCACTGGGATCAATCCCAATAGGAGTGGAGTCAAG 4200
 Db 4141 GTTGGGGGCTTACTTGTCTTAAAGGCACTGGGATCAATCCCAATAGGAGTGGAGTCAAG 4200
 Qy 4201 GACTGTGACGACCGGGGGCGCCCATCAGTACTTCCACATATGGCAATTTCTTCCCGCATG 4260
 Db 4201 GACTGTGACGACCGGGGGCGCCCATCAGTACTTCCACATATGGCAATTTCTTCCCGCATG 4260
 Qy 4261 GGGCTGCGGGGGGCGGCTGAGCATCATATGATGATGATGATGATGATGATGATGATGATGATG 4320
 Db 4261 GGGCTGCGGGGGGCGGCTGAGCATCATATGATGATGATGATGATGATGATGATGATGATGATG 4320
 Qy 4321 TACCAACATCTTGGGATGCGAAGCTCTTGTATCAAGCAGAGACAGCTGGGTCAGACT 4380
 Db 4321 TACCAACATCTTGGGATGCGAAGCTCTTGTATCAAGCAGAGACAGCTGGGTCAGACT 4380
 Qy 4381 AACTGTGCTGCTCAGACTTACGCTTGGGCTGAGTCAACCCCGCATCCCAATAGATA 4440
 Db 4381 AACTGTGCTGCTCAGACTTACGCTTGGGCTGAGTCAACCCCGCATCCCAATAGATA 4440
 Qy 4441 GAGGTGGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4500
 Db 4441 GAGGTGGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4500
 Qy 4501 TTATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4560
 Db 4501 TTATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4560
 Qy 4561 CCGCGGGGCTTCTGGGGGTATGGGCTTGAATCTGAGTGGGCTTGAAGGAGGAGGAGGAGGAGGAG 4620
 Db 4561 CCGCGGGGCTTCTGGGGGTATGGGCTTGAATCTGAGTGGGCTTGAAGGAGGAGGAGGAGGAGGAG 4620
 Qy 4621 CTCGGTAAACCACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680
 Db 4621 CTCGGTAAACCACTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4680
 Qy 4681 GTATATCTGGGAGCTTTGATCTCGGTGATGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4740
 Db 4681 GTATATCTGGGAGCTTTGATCTCGGTGATGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4740

QY	8041	CCTCTGGAGACTCAGAAACCAATTTCCCAAAACCAATTTAGTCCAAAATAGAGTGTT	8100
DB	8041	CCTCTGGAGACTCAGAAACCAATTTCCCAAAACCAATTTAGTCCAAAATAGAGTGTT	8100
QY	8101	CTCGTGGACCCCAACCAAGGGGGGCAAGAAAGACGCTCGCCTTATCGTTTACCCCTGACCT	8160
DB	8101	CTCGTGGACCCCAACCAAGGGGGGCAAGAAAGACGCTCGCCTTATCGTTTACCCCTGACCT	8160
QY	8161	CGGCGTTCAGGGTCTGCGAGAGATGSCCTTTATGACATATACAAAAAATCTTCTCAGGC	8220
DB	8161	CGGCGTTCAGGGTCTGCGAGAGATGSCCTTTATGACATATACAAAAAATCTTCTCAGGC	8220
QY	8221	GGTCATGGGGGCTCTTATTGATATCCAGTATTCGCCGCTCAGCGGGTAGAGTTCTCTTT	8280
DB	8221	GGTCATGGGGGCTCTTATTGATATTCAGTATTCGCCGCTCAGCGGGTAGAGTTCTCTTT	8280
QY	8281	GAAGCATGGGGCGGAAAGAAAGACCTATGSGTTTTTCGTATGATACCCGATGCTTTGA	8340
DB	8281	GAAGCATGGGGCGGAAAGAAAGACCTATGSGTTTTTCGTATGATACCCGATGCTTTGA	8340
QY	8341	CTCAACCGTCACTGAGAGAGACATCAGGACTGAGAGTCCATATATCGGSCCTGCTCTTT	8400
DB	8341	CTCAACCGTCACTGAGAGAGACATCAGGACTGAGAGTCCATATATCGGSCCTGCTCTTT	8400
QY	8401	GCCCGAGGGCCCACTGCCATACACTCGCTAACTGAGAGACTTTACGTCGGAGGSGCC	8460
DB	8401	GCCCGAGGGCCCACTGCCATACACTCGCTAACTGAGAGACTTTACGTCGGAGGSGCC	8460
QY	8461	TATGTTCAACAGCAGGGCCAAACCTGGGGTACAGGGTTTCGCGCCGAGCGGGTGCT	8520
DB	8461	TATGTTCAACAGCAGGGCCAAACCTGGGGTACAGGGTTTCGCGCCGAGCGGGTGCT	8520
QY	8521	CACCATAGCATGGGAAACCATCACATGCTACGTGAAGCCTTAGCGGCTTGTAAAGC	8580
DB	8521	CACCATAGCATGGGAAACCATCACATGCTACGTGAAGCCTTAGCGGCTTGTAAAGC	8580
QY	8581	TGCAGGGATTAATCGCGCCCACAATCTGGTATGCGGGGATGACTTGGTTGTCACTCAGA	8640
DB	8581	TGCAGGGATTAATCGCGCCCACAATCTGGTATGCGGGGATGACTTGGTTGTCACTCAGA	8640
QY	8641	AAGCCAGGGACCGAGAGGACAGCGGAACCTCAGAGCCTTACGAGGCTATGACCAG	8700
DB	8641	AAGCCAGGGACCGAGAGGACAGCGGAACCTCAGAGCCTTACGAGGCTATGACCAG	8700
QY	8701	GTAATTCCTGCCCCTCTGGTGACCCCGCCAGACGGAGTATGATCTGGAGTGATACATC	8760
DB	8701	GTAATTCCTGCCCCTCTGGTGACCCCGCCAGACGGAGTATGATCTGGAGTGATACATC	8760
QY	8761	TTGCTCTCAAAATGTTCTGTGGGCGTGGGCGCCACAAGCGCGCGCAGATATCTAGC	8820
DB	8761	TTGCTCTCAAAATGTTCTGTGGGCGTGGGCGCCACAAGCGCGCGCAGATATCTAGC	8820
QY	8821	CAGAGACCTTACACTCCAAATCGCGCGGGCTGCCCTGGGAAACAGTTAGACACTCCCTGT	8880
DB	8821	CAGAGACCTTACACTCCAAATCGCGCGGGCTGCCCTGGGAAACAGTTAGACACTCCCTGT	8880
QY	8881	CAATTATGGCTGGGAAACATCATTCAGTACGCCCGCACCATATGGGCTCGCATGGTCT	8940
DB	8881	CAATTATGGCTGGGAAACATCATTCAGTACGCCCGCACCATATGGGCTCGCATGGTCT	8940
QY	8941	GATGACACTTCTTCTCCATCTCATGGGCTCAGACACGCTGGACAGAACTCAACTT	9000
DB	8941	GATGACACTTCTTCTCCATCTCATGGGCTCAGACACGCTGGACAGAACTCAACTT	9000
QY	9001	TGAGATGATACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCGAGCTATAATTGAAAG	9060
DB	9001	TGAGATGATACGAGCGGTGTACTCCGTGAGTCCCTTGGACCTCCGAGCTATAATTGAAAG	9060
QY	9061	GTTACATGGGCTTGACGCTTTTCTCTGCAACATACACTCCCCACGAACTGACACGGGT	9120
DB	9061	GTTACATGGGCTTGACGCTTTTCTCTGCAACATACACTCCCCACGAACTGACACGGGT	9120

WPI; 1993-087166/11.
N-Psda; AAR3536.

Poly nucleotide(s), polypeptide(s) and antibodies of NANBH virus -
useful for detecting NANBH, as a vaccine and for screening blood
samples

Claim 2 ; Page 27-32; 93pp: English.

RNA was isolated from the plasma of human patients positive for
NANBH virus (strain HC-J6) and was subjected to reverse transcription
to produce cDNA. The resulting cDNA was amplified by PCR, and
nucleic acid sequences determined by analysis of both clones from the
cDNA library and clones obtd. by PCR amplification (36 clones in total).
The NANBH HC-J6 genome was found to contain an open reading frame
encoding a polypeptide precursor of 3033 amino acid residues.
See also AAQ318172-221.
(Updated on 25-MAR-2003 to correct PN field.)

Sequence 9589 BP: 1968 A; 2920 C; 2635 G; 2166 T; 0 other;

Query Match	92.4%	Score 8970.4	DB 14	Length 9589
Best Local Similarity	96.0%	Pred. No. 0		
Matches 9202	Conservative 0	Mismatches 386	Indels 0	Gaps 0
QY	1	ACCCGCCCTTAATAGGGGGGCGACACTCCGGCCATGAATCACTCCCTGTGTGAGGAATCACTACTGT	60	
DB	1	ACCCGCCCTTAATAGGGGGGCGACACTCCGGCCATGAACCACTCCCTGTGTGAGGAATCACTACTGT	60	
QY	61	CTTCAGCGCACAAAGCGCTCTAGCCCATGGCGTTAGTATGAGTGTCTGTACAGCCTCCAGAGCCC	120	
DB	61	CTTCAGCGCAAAAGCGCTCTAGCCCATGGCGTTAGTATGAGTGTCTGTACAGCCTCCAGAGCCC	120	
QY	121	CCCCCTCCGGGAGAGCCATAGTGTCTCTCGGAAACCGGTGAGTATACCGGAATTTCCGGG	180	
DB	121	CCCCCTCCGGGAGAGCCATAGTGTGTCTCGGAAACCGGTGAGTATACCGGAATTTCCGGG	180	
QY	181	AAGACTGGGTCTCTTTCTTGATATAAACCCACTCTATGTCCCGGCCATATTTGGGGGTGCCCCCG	240	
DB	181	AAGACTGGGTCTCTTTCTTGATATAAACCCACTCTATGTCCCGGCCATATTTGGGGGTGCCCCCG	240	
QY	241	CAAGACTGTCTAGCGGAGTAGCGTTGGGTTTGCAGAAAGGCTCTGTGTACTCTGCTCATATAGG	300	
DB	241	CAAGACTGTCTAGCGAGTAGCGTTGGGTTTGCAGAAAGGCTCTGTGTACTCTGCTCATATAGG	300	
QY	301	TGCTTGCGAGTGTCCCGGGAGGTCTCGTATAGACCGGTGCAACATGTAGACACAATACTCTAAACC	360	
DB	301	TGCTTGCGAGTGTCCCGGGAGGTCTCGTATAGACCGGTGCAACATGTAGACACAATACTCTAAACC	360	
QY	361	TCAAAGAAAAACCAAAGAAAAACCAACACCGTCGCCCAACAAGACGTTAAGTTTCCGGCGG	420	
DB	361	TCAAAGAAAANACCAAAGAAAAACCAACACCGTCGCCCAACAAGAGGTTAAGTTTCCGGCGG	420	
QY	421	CGCGCAGATCGTTGGGGAGTATACTTGTTCGGCGCGCAGGGGGCCCAAGTTGGGTGTGCG	480	
DB	421	CGCGCAGATCGTTGGGGAGTATACTTGTGTTCGGCGCGCAGGGGGCCCAAGTTGGGTGTGCG	480	
QY	481	CGCGACAAGGAAGAATTTCGGAGCGGTCCGAGCCAAGTGGAGGGCGCCAGCCATCCCTTAA	540	
DB	481	CGCGACAAGGAAGAATTTCGGAGCGGTCCGAGCCAAGTGGAGGGCGCCAGCCATCCCTTAA	540	
QY	541	AGATGGCGCTTCACTTGGCAAAATCTTGGGGGAAAAACAGGATACCCCTGGCCCCCTATACGG	600	
DB	541	GGATGGCGCTTCACTTGGCAAAATCTTGGGGGAAAAACAGGATACCCCTGGCCCCCTATACGG	600	
QY	601	GAATGAGGGAAGTCCGGTGGCGAGGATGCGTCTCTGTCTCCCGAGAGTTCCTCGCTCTTG	660	
DB	601	GAATGAGGGAAGTCCGGTGGCGAGGATGCGTCTCTGTCTCCCGAGAGTTCCTCGCTCTTG	660	
QY	661	GGGCCCCCAATGAACCCCCCGGATAGGTGTGCGGAAACGTGGGGTAAGGTTCATCGATACCTTAAC	720	
DB	661	GGGCCCCCAATGAACCCCCCGGATAGGTGTGCGGAAACGTGGGGTAAGGTTCATCGAACCTTAAC	720	

RESULT 4	
AAQ38218	
ID	AAQ38218 standard; cDNA to mRNA; 9589 BP.
XX	
XX	AAQ38218;
XX	AC
XX	
DT	25-MAR-2003 (updated)
DT	01-JUL-1993 (first entry)
XX	
DE	NANBH virus strain HC-J6 genome cDNA sequence.
XX	
XX	Non A non B hepatitis virus; amplification; HC-J1; HC-J8; plasma; ss.
XX	
OS	Non A, non B hepatitis virus strain HC-J6.
XX	
Key	Location/Qualifiers
PH	341...9442
FT	CDs
FT	/*tag= a
XX	
PN	EP532167-A2.
XX	
PD	17-MAR-1993.
XX	
XX	30-JUL-1992; 92EP-0306952.
XX	
PP	09-AUG-1991; 91JP-0287402.
PR	05-DEC-1991; 91JP-0360441.
XX	
PA	(IMMO) IMMUNO JAPAN INC.
XX	
PI	Nakamura T, Okamoto H;
XX	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2004, 06:04:40 ; Search time 13743 Seconds
(without alignments)
11773.505 Million cell updates/sec

Title: US-09-980-559-1

Perfect score: 9711

Sequence: 1 accgcgccctaataggggcg.....ggtctctcgcagatcatgt 9711

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_hg_hum.*
31: em_hg_inv.*
32: em_hg_other.*
33: em_hg_mus.*
34: em_hg_pln.*
35: em_hg_rdi.*
36: em_hg_mam.*
37: em_hg_vrt.*
38: em_sy.*
39: em_hgo_hum.*
40: em_hgo_mus.*
41: em_hgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9711	100.0	9711	6	AX057086 Sequence
2	9711	100.0	9711	6	AX057317 Sequence
3	9711	100.0	9711	6	AX057395 Sequence
4	9711	100.0	9711	14	AF177036 Sequence
5	8973.6	92.4	9589	14	HPCPOLP
6	8970.4	92.4	9589	6	E07361
7	8970.4	92.4	9589	6	E07362
8	8970.4	92.4	9589	6	E12861
9	8507.4	87.6	9659	14	AB047641
10	8425.8	86.8	9700	14	AF169005
11	8370.2	86.2	9693	14	AF169003
12	8367.8	86.2	9653	14	AF169004
13	8364.2	86.1	9683	14	AB047645
14	8349	86.0	9691	14	AB047644
15	8329.4	85.8	9661	14	AF169002
16	8273.4	85.2	9674	14	AB047640
17	8233.4	84.8	9678	14	AB047642
18	8210.2	84.5	9416	14	AF238481
19	8186.8	84.3	9416	14	AF238482
20	8169.2	84.1	9416	14	AF238483
21	8156.4	84.0	9416	14	AF238484
22	8126.4	83.7	9416	14	AF238485
23	8083.4	83.2	9641	14	AB047643
24	8020.6	82.6	9678	6	BD160781
25	8020.6	82.6	9678	14	AB047639
26	6479.2	66.7	9513	14	D50409
27	6383.6	65.7	9488	14	AB031663
28	6048	62.3	9654	14	AB030907
29	6031.4	62.1	9511	6	E07468
30	6031.4	62.1	9511	6	E12865
31	6027	62.1	9511	6	E07467
32	5933.8	61.7	9511	14	HPC08G
33	5936	61.1	9416	14	AF238486
34	5816.8	59.9	9611	14	AF177037
35	5800.8	59.7	9611	6	AX057088
36	5800.8	59.7	9611	6	AX057092
37	5800.8	59.7	9611	14	AF177039
38	5692	58.6	9611	6	AX057090
39	5692	58.6	9611	14	AF177038
40	5676	58.4	9611	6	AX057094
41	5676	58.4	9611	14	AF177040
42	4625.6	47.6	9646	6	AB110828
43	4625.6	47.6	9646	6	BD069982
44	4625.6	47.6	9646	14	AF009606
45	4620.8	47.6	9646	12	AF387806

ALIGNMENTS

RESULT 1	AX057086	Sequence 1 from Patent WO0075338.	9711 bp	DNA	linear	PAT 17-JAN-2001
LOCUS	AX057086	Sequence 1 from Patent WO0075338.				
DEFINITION	AX057086					
ACCESSION	AX057086					
VERSION	AX057086.1	GI:12309919				
KEYWORDS						
SOURCE	Hepatitis C virus					
ORGANISM	Hepatitis C virus					
	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;					
REFERENCE	1					
AUTHORS	Yanagi M., Bukh J., Emerson S.U. and Purcell R.H.					
TITLE	Cloned genome of infectious hepatitis C virus of genotype 2a and uses thereof					

Pred. No. is the number of results predicted by chance to have a

ORGANISM Hepatitis C virus
 Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 Hepacivirus
 1 (bases 1 to 9589)
 Okamoto, H., Okada, S., Sugiyama, Y., Kurai, K., Iizuka, H., Machida, A.,
 Miyakawa, Y., and Mayumi, M.
 Nucleotide sequence of the genomic RNA of hepatitis C virus
 isolated from a human carrier: comparison with reported isolates
 for conserved and divergent regions
 J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
 92044440
 JOURNAL MEDLINE
 PUBMED
 REFERENCE
 1658196
 2 (sites)
 Han, J. H. and Houghton, M.
 Group specific sequences and conserved secondary structures at the
 3' end of HCV genome and its implication for viral replication
 Nucleic Acids Res. 20 (13), 3520 (1992)
 92335016
 JOURNAL MEDLINE
 PUBMED
 REFERENCE
 1321416
 3 (sites)
 Hotta, H., Doi, H., Hayashi, T., Purwanta, M., Soemarto, W.,
 Mizokami, M., Ohba, K. and Honma, M.
 Analysis of the core and E1 envelope region sequences of a novel
 variant of hepatitis C virus obtained in Indonesia
 Arch. Virol. 136 (1-2), 53-62 (1994)
 94370990
 JOURNAL MEDLINE
 PUBMED
 REFERENCE
 7545932
 These data kindly submitted in computer readable form by: Hiroaki
 Okamoto
 Immunology Division
 Jichi Medical School
 Kawachi-gun
 Tochigi-Ken 329-04
 Japan
 Phone:
 Fax:
 0285-44-2111 x3334
 0285-44-1557.
 Location/Qualifiers
 1. 9589
 /organism="Hepatitis C virus"
 /mol_type="genomic RNA"
 /isolate="HC-J6"
 /db_xref="taxon:11103"
 341..9442
 /codon_start=1
 /product="polyprotein"
 /protein_id="BA00792.1"
 /db_xref="GI:221651"
 /translation="MSTNPKPQKTKNTNRPDQVKFGGQIVGGVYLLPRGPRL
 GVRATKTSRSPRRQPIPDORRSTGKSGKPGYPMPLYGNEGLWAGWLLSPRG
 SRPSMGNDPRSRNVKVIDTLTCGADLMGYIPVVGAPLGGVAAALAHGVRVLED
 GYNFATNLPCCSFSIFLLALSCITTPSAAEVKNISGTVMYNDNDNSITWLOQA
 AVLVPGVPCSKVGTNRKCNIPVSNVAVQPGALTOGLRTHIDMVMSATLCSALY
 VGLCCGVMAAQMTIVSPQHHMVFQDCNCSLYPGTITGRHAWDMRWNSFATNLL
 AVAMRVPVSDIIGAGKGMFGLAYFSMQQANAKVYVILLAAQVDHQHTVGGST
 ANHARTLCMFLGARQKLIQINTNGSHINRINALNCNDELSTFGLASLFTYHNS
 CPERMSACRSLEAFVWGALQYEDNVTPEMRPYCYHYPQCGVVSASSVGVYK
 YCTFSPVWGTDRLGATYTWGENETDVLNSTRPFGSNFGCTWNSGTGTYKTC
 GAPPCIRDFNASHMDDLCPDTCRPHDPTTYIKCGSPWLTPLCLIDYPRVLMHPC
 TWNYTFKIRMYGVGHRHTAACNFTGRDCLNEDRDSQLSPLLHSITTEMAILPCT
 YSDPLSTGLLHLHQNIVQVFMGLSPALTYIVRWKVVLLFLLADARVCACLW
 MLTLQGAALAEKLVLRASAAASCNGLYFVIFVAAWVYIKGRVPLATYVLTGLM
 SFGLLIALPQOAVAYDASVHQIQAALLVLTFLTPGKTLISRLFLMCLYLLTL
 AEMVQENAPPMQVGGSDGIIMAVAIKFCPGVVDITKMLAVLAPVLLKGLTRP
 YFVRHALLRMEKTVHLAGGRYQWVLLALGRWGTLYDHLTPMSWANGRLDLA
 VAEPIIFSPMEKTVHLAGGRYQWVLLALGRWGTLYDHLTPMSWANGRLDLA
 APITAYAGTGLGLGTIVSWNTGDKTEQAGEIQLYSTVQSLGTISGVLMYVING
 AGNKTLAGSGRPTQMTSSAGDLYGMPSPGKLSLEPCTCGAVDLIVTRNADYIPA
 RRGKRGALLSPFLSTLKGSGGVLCPGRHAGVFRAAVCSRGVAKSIDFIPVET
 LQIVYRPTFSNSTRPAPVQTYGVYHAPTSKSTKVPVAYAAQGYKVLVNPVS
 ATGLGFAVLSKAGHNINRTVTVTGAPITYSTYKFLADGGCAGGAYDIICD
 ECHAVDSITLIGTVLQDQTAGVRLTVLATATPPGVTTPPHNIBEVAGSGEIP
 FYGRAIPUSYINGGSHLIFCHSKKCDLALRGNLNAVATYRGLDVSIVTPQGVY
 VVATDALMTGTGDPDSVIDONVAVTVQVVDVFLDPTFTITQTVPQDAVRSORRRG

FEATURES
 source

CDS

BASE COUNT 1970 a 2820 c 2633 g 2166 t
 ORIGIN
 Query Match 92.4%; Score 8973.6; DB 14; Length 9589;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 9204; Conservative 0; Mismatches 384; Indels 0; Gaps 0;
 Qy 1 ACCCGCCCTAATAGCGGGACATCGCCCATGAACTACCTCCCTGTGAGGAACCTACTGT 60
 Db 1 ACCCGCCCTAATAGCGGGACATCGCCCATGAACTACCTCCCTGTGAGGAACCTACTGT 60
 Qy 61 CTTTCAGCAGAAAGCGTCTAGCCATGCGGTAGTAGTGTCGTACACGCTCCAGGCCC 120
 Db 61 CTTTCAGCAGAAAGCGTCTAGCCATGCGGTAGTAGTGTCGTACAGCCTCCAGGCCC 120
 Qy 121 CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGG 180
 Db 121 CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGG 180
 Qy 181 AAGACTGGGTCTCTTCTTGGATAAACCCACTATATCCCGGCCNTTGGCGGTGCGCCG 240
 Db 181 AAGACTGGGTCTCTTCTTGGATAAACCCACTATATCCCGGTCACTTGGCGGTGCGCCG 240
 Qy 241 CAAGACTCTAGCCGAGTAGCGTGGGTTCGGAAGCGCTCTGTGTGCTAGTCCCTGATAGG 300
 Db 241 CAAGACTCTAGCCGAGTAGCGTGGGTTCGGAAGCGCTCTGTGTGCTAGTCCCTGATAGG 300
 Qy 301 TCGTTGCGAGTGCCCGGAGGTCTCTGTAGACGCTGCACCATGAGCACAATCTCTAAACC 360
 Db 301 TCGTTGCGAGTGCCCGGAGGTCTCTGTAGACGCTGCACCATGAGCACAATCTCTAAACC 360
 Qy 361 TCARAGAAAAACCAAAAGAAACCAACCTCCGCCACAGAGCTTAAGTTTCGCGGGG 420
 Db 361 TCARAGAAAAACCAAAAGAAACCAACCTCCGCCACAGAGCTTAAGTTTCGCGGGG 420
 Qy 421 CGGCCAGATCTGTGGCGAGTATACCTTGTTCGCGGCGAGGGGCCCGAGTTGGGTGCGG 480
 Db 421 CGGCCAGATCTGTGGCGAGTATACCTTGTTCGCGGCGAGGGGCCCGAGTTGGGTGCGG 480
 Qy 481 CGCGCAAGAGGAGACTTCGAGCGGTCCAGGCCACCTGGAAGGCGCCAGCCCATCTCTAA 540
 Db 481 CGCGCAAGAGGAGACTTCGAGCGGTCCAGGCCACCTGGAAGGCGCCAGCCCATCTCTAA 540
 Qy 541 AGATCGGCGCTCCATCGCAAACTCTGGGGAACACAGGATACCTCGCCCGCTTATAGG 600
 Db 541 AGATCGGCGCTCCATCGCAAACTCTGGGGAACACAGGATACCTCGCCCGCTTATAGG 600

W 00075

1603

Db 9221 GCGTTTGGCGGATATCTTCACTGGCGGTGAGCCAGCTCAACTCACTCA 9280
 Qy 2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTrpPheThrValGlyAlaGlyGly 3000
 Db 9281 TTCCCGGAGCGCGCTCTGATTTATTCAGCTGGTTCACTGTCCGCGCGCGGGGCG 9340
 Qy 3001 AspIleYrHisSerValSerArgAlaArgProArgLeuLeuPheGlyLeuLeuLeu 3020
 Db 9341 GACATTATACAGCGGTGCGGTGCGCGCGCGCTTATTACTCTTGGCTACTCTTA 9400
 Qy 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 Db 9401 CTTTTTGAAGGGAAGGCTTTCTACTTCCCGCTCGG 9439

RESULT 8

LOCUS 112861
 DEFINITION Sequence 2 from patent US 5428145.
 ACCESSION 112861
 VERSION 112861.1 GI:910242
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFS 1 (bases 1 to 9589)
 AUTHORS Okamoto,H. and Nakamura,T.
 TITLE Non-A, non-B, hepatitis virus genome, polynucleotides,
 polypeptides, antigen, antibody and detection systems.
 JOURNAL Patent: US 5428145-A 2 27-JUN-1995;
 FEATURES
 source 1..9589
 /organism="unknown"

BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.15% Indels: 0
 DB: 6 Gaps: 0

US-09-980-559-2 (1-3033) x 112861 (1-9589)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 341 ATGAGCAACAATCTTAACCTCAAGAGAAACCAAGAAACCAACCAACCGTGGCCCA 400
 Qy 21 AspValLysPheProGlyGlyGlnIleValGlyValGlyValGlyValGlyValGly 40
 Db 401 GAGTTAAGTTTCGGCGCGCGCGCAGATCGTTGGCGGAGTATCTTGTTCGGCGCAG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60
 Db 461 GCGCCGAGGTGGGTGGCGCGCGACAGAGAGACTTCGAGCGGTCCGAGCCACGTTGA 520
 Qy 61 ArgArgGlnProIleProLysAspArgSerThrGlyLysSerTrpGlyLysProGly 80
 Db 521 AGCGCGCAGCCCATCCCTAAGGATCGCGCTCCACTGGCAATCTCGGGAACCAAGCA 580
 Qy 81 TyrProTrpProLeuTyrGlyAsnGluGlyLeuGlyTrpAlaGlyTrpLeuLeuSerPro 100
 Db 581 TACCCTTGGCCCTATACGGGATGAGGACTCGGTGGCGGAGATGGCTCTGTCTCCCC 640
 Qy 101 ArgGlySerArgProSerTrpGlyProAsnAspProArgHisArgSerArgAsnValGly 120
 Db 641 CGAGGTTCCTGCTCTTGGCGCCCAATGACCCCGCATAGGTCCCGCAACGTTGGGT 700
 Qy 121 LysValIleAspThrLeuThrCysGlyPheAlaAspLeuMetGlyTyrIleProValVal 140
 Db 701 AAGGTCATCATGATCCCTAAGCTGGCGCTTTGGCGACTCATGGGGTACATCCCTGTGTA 760

Qy 141 GlyAlaProLeuGlyGlyValAlaAlaArgAlaLeuAlaHisGlyValArgValLeuGluAsp 160
 Db 761 GCGCCCGCGCTCGCGCGCGCGTCCAGAGCTCTCGCGCATGGCGTGGAGAGTCTCTGGAG 820
 Qy 161 GlyValAsnPheAlaThrGlyAsnLeuProGlyCysSerPheSerIlePheLeuLeuAla 180
 Db 821 GGGGTAAATTTTCAACAGGAACCTTACCCTGGTCTCTCTTTCTATCTCTCTCTCTG 880
 Qy 181 LeuLeuSerCysIleThrThrProValSerAlaAlaGluValLysAsnIleSerThrGly 200
 Db 881 CTGCTGTCTCTGATCAACACCCCGGTCTCCGCTCGGAGTGAAGACATCATGATCCGCG 940
 Qy 201 TyrMetValThrAsnAspCysThrAsnAspSerIleThrTrpGlnLeuGlnAlaAlaVal 220
 Db 941 TACATGTGACCAACGACCTGACCAATATGATAGCATTTACTTGGCACTCCAGGCTGCT 1000
 Qy 221 LeuHisValProGlyCysValProCysGluLysValGlyAsnAlaSerGlnCysTrpIle 240
 Db 1001 CTCCACGCTCCCGGTGCGTCCCGTGGAGAAAGTGGGGAATACATCTCGTGTCTGATA 1060
 Qy 241 ProValSerProAsnValAlaValGlnArgProGlyAlaLeuThrGlnGlyLeuArgThr 260
 Db 1061 CCGGTCTCACCGAATGTGCGGTGCGAGCCCGCGGCTTCCAGCGAGGCTTACGAGCG 1120
 Qy 261 HisIleAspMetValValMetSerAlaThrLeuCysSerAlaLeuTyrValGlyAspLeu 280
 Db 1121 CACATTGACATGGTTGTGATGTCCGCGAGCTCTGCTCTCTCTCTCTCTCTCTCTCT 1180
 Qy 281 CysGlyGlyValMetLeuAlaAlaGlnMetPheIleValSerProGlnHisHisTrpPhe 300
 Db 1181 TCGGTGGGTGATGCTTGCAGCCAGATGTTCTATGTTCTGCCACAGCACCACTGGTTT 1240
 Qy 301 ValGlnAspCysAsnCysSerIleTyrProGlyThrIleThrGlyHisAspMetAlaTrp 320
 Db 1241 GTGCAAGACTGCAATTTGCTTCCATCTACCTGGTACCATCACTGCGACACCGTGGG 1300
 Qy 321 AspMetMetAsnTrpSerProThrAlaThrMetIleLeuAlaTyrAlaMetArgVal 340
 Db 1301 GACATGATGATGACCTGGTGGCGCGCTACCATGATCTTGGCGTGGCGTGGCGGTC 1360
 Qy 341 ProGluValIleIleAspIleIleSerGlyAlaHisTrpGlyValMetPheGlyLeuAla 360
 Db 1361 CCGAGGTCTATCATAGACATCATTTGGCGGGCTCATTTGGGGCGTCTCATTTGGGCT 1420
 Qy 361 TyrPheSerMetGlnGlyAlaTrpAlaLysValValIleLeuLeuLeuAlaAlaGly 380
 Db 1421 TACTTCTCTATGCGAGCGGTGGCAAAAGTGGTTGTGTCATTTCTTTTGTCTGGCG 1480
 Qy 381 ValAspAlaArgThrHisThrValGlyGlySerAlaAlaGlnThrThrGlyArgLeuThr 400
 Db 1481 GTGGACCGCAACCCATACCGTGGGGTCTACCGCGCATTAACGCCAGACCCCTCACC 1540
 Qy 401 SerLeuPheAspMetGlyProArgGlnLysIleGlnLeuValAsnThrAsnGlySerTrp 420
 Db 1541 GGCATGTTCTCTCTGGTCCAGGCGAGAAATCCAGCTCATCAACACCAATGGCAGTTG 1600
 Qy 421 HisIleAsnArgThrAlaLeuAsnCysAsnAspSerLeuHisThrGlyPheIleAlaSer 440
 Db 1601 CACATCAACCGCACCCCTGAACTCTTTGTCACACTCTTTGTCACACCGGCTCTCTG 1660
 Qy 441 LeuPheTyrThrHisSerPheAsnSerSerGlyCysProGluArgMetSerAlaCysArg 460
 Db 1661 CTGTTCTACCCACAGCTTCACTCTGAGGATGTCGAGAACGATGTCGCTCTGCGC 1720
 Qy 461 SerIleGluAlaPheArgValGlyTrpGlyAlaLeuGlnTyrGluAspAsnValThrAsn 480
 Db 1721 AGTATCGAGGCTCTTGGGTGGGATGGGGCGCTTTACAAATATAGGACAAATGTCACCA 1780
 Qy 481 ProGluAspMetArgProTyrCysTrpHisTyrProProArgGlnCysGlyValValSer 500
 Db 1781 CAGAGGATATGAGACCGTATGCTGGCACTACCCACCAAGACAGTGTGCTGTGTCTCC 1840

8141 CTTATCGTTTACCTGACCTCGGCTCAGGCTTCGCGAGAAATGCGCCCTTATGATATC 8200
 2621 ThrGlnLeuProGlnAlaValMetGlyAlaSerTyrGlyPheGlnTyrSerProAla 2640
 8201 ACAAAAGCTTCCTCAGCGGTGATGGGGCTCTTATGGAATTCAGTACTCCCGGCT 8260
 2641 GlnArgValGluPheLeuLeuLysAlaTTPAlaGluLysLysAspProMetGlyPheSer 2660
 8261 CAGCGGTGAGTTCCTCTTGAAGGATGGCGGGAAGAACACCTATGGTTCG 8320
 2661 TyrAspThrArgCysPheAspSerThrValThrGluArgPheLeuThrGluGluSer 2680
 8321 TATGATACCGATGCTTTCATCAACCTCACTGAGAGACATCAGGACTCAGAGATCC 8380
 2681 IleTyrArgAlaCysSerLeuProGluAlaHisThrAlaIleHisSerLeuThrGlu 2700
 8381 ATATATCGGGCTTGTTCCTGCGGAGGAGCCACACTGCCATACACTCACTGCTGAG 8440
 2701 ArgLeuTyrValGlyProMetPheAsnSerLysGlyGlnThrCysGlyTyrArgArg 2720
 8441 AGACTTTAGCTGGAGGGCCCATGTTCAACGACGAGGCCAGACCTGCGGGTACAGGCT 8500
 2721 CysArgAlaSerGlyValLeuThrThrSerMetGlyAsnThrIleThrCysTyrValLys 2740
 8501 TGGCGGCGCAGCGGGTCTTACCACTAGCATGGGACACCATCATGCTATGTGAA 8560
 2741 AlaLeuAlaCysGlyAlaAlaGlyIleIleAlaProThrMetLeuValCysGlyAsp 2760
 8561 GCCTTAGCGGCTGTAAAGCTGCAAGGATTAATTCGCGCCCAATGCTGTATGCGGCGAT 8620
 2761 AspLeuValIleSerGluSerGlnGlyThrGluAspGluArgAsnLeuArgAla 2780
 8621 GACTTGTGTCTCATCTCAGAGAGCGAGGGAGCCAGGAGGAGGAGCGGAACCTGAGAGCC 8680
 2781 PheThrGluIleMetThrArgTyrSerAlaProGlyAspProProArgProGluTyr 2800
 8681 TTCAGGAGGCTATGACCACTATCTGCGCCCTCTGCTGAGCCGCCAGAGCGGAAAT 8740
 2801 AspLeuGluLeuIleThrSerCysSerSerAsnValSerValAlaLeuGlyProGlnGly 2820
 8741 GACTTGGAGCTGATTAACATCTTCTCTCTCAAAATGCTGTGGCGGTGGGCCCAACAGGC 8800
 2821 ArgArgArgTyrTyrLeuThrArgAspProThrThrProIleAlaArgAlaAlaTrpGlu 2840
 8801 CGCGGACAGATACCTCAGCAGAGACCTACCACTCCAATCGCGGCTGCGGCTGCGCA 8860
 2841 ThrValArgHisSerProValAsnSerTyrLeuGlyAsnIleIleGlnTyrAlaProThr 2860
 8861 ACAGTTAGACACTCCCTGTCTCATTTCTGCTAGGAAACATCATCAGTACGCCCCCAACC 8920
 2861 IleTyrAlaArgMetValLeuMetThrHisPheSerIleLeuMetAlaGlnAspThr 2880
 8921 ATATGGCTCGCATGCTCTCATGACACACTTCTTCTCAATCTCTATGCGCCCAAGTACT 8980
 2881 LeuAspGlnAsnLeuAsnPheGluMetTyrGlyAlaValTyrSerValSerProLeuAsp 2900
 8981 CTGGACAGAACCTCAACTTTCAGATGTAGCGAGCGGTGACTCTCGTGAATCCCTTGGAC 9040
 2901 LeuProAlaIleIleGluArgLeuHisGlyLeuAspAlaPheSerLeuHisThrTyrThr 2920
 9041 CTCCGAGCCATAATTGAAGGTACACGGGCTTGACGCTTCTCTCTGACACATACACT 9100
 2921 ProHisGluLeuThrArgValAlaSerAlaLeuArgLysLeuGlyAlaProProLeuArg 2940
 9101 CCCAGCAACTGACACGGGTGGCTTCAGGCCCTCAGAAATCTTGGGGCGGCCACCTCAGA 9160
 2941 AlaTrpLysSerArgAlaArgAlaValArgAlaSerLeuIleSerArgGlyValArgAla 2960
 9161 GGTGGAGAGCGCGGACGTCGAGTCAAGGCGCTCCCTCATCTCCCGTGGGGGAGAGCG 9220
 2961 AlaValCysGlyArgTyrLeuPheAsnTTPAlaValIleThrLysLeuLeuLeuThrPro 2980
 9221 GCGGTTGGCGCGATATCTCTTCACTTGGCGGCTGAGAGCCAGACTCAAACTCACTCA 9280

2981 LeuProGluAlaArgLeuLeuAspLeuSerSerTyrPheThrValGlyValaGlyGly 3000
 9281 TTGCGGAGAGCGGCTCTCTGATTTTCCAGCTGGTTCACTGTCGCGCGCGCGGCG 9340
 3001 AspileTyrHisSerValSerArgAlaAspProArgLeuLeuPheGlyLeuLeu 3020
 9341 GACATTTATCACAGCGTGTGCGTGCAGACCCCTTATTAATCTCTTGGCTACTCTCA 9400
 3021 LeuPheValGlyValGlyLeuPheLeuProAlaArg 3033
 9401 CTTTTCGTAGGGTAGCGCTTTTCTTACTCCCGCGCTGG 9439
 RESULT 7
 E07362
 LOCUS CDNA of Hepatitis non-A non-B virus. linear PAT 29-SEP-1997
 DEFINITION
 E07362
 VERSION E07362.1 GI:2175501
 KEYWORDS JP 1994121689-A/2.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1. (bases 1 to 9589)
 AUTHORS Okamoto, H. and Nakamura, T.
 TITLE NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
 ANTIGEN AND ANTIBODY DETECTION SYSTEM
 JOURNAL Patent: JP 1994121689-A 2 06-MAY-1994.
 COMMENT
 OS Hepatitis non-A non-B virus
 PN JP 1994121689-A/2
 PD 06-MAY-1994
 PF 09-AUG-1991 JP 1991287402
 PI OKAMOTO HIROAKI, NAKAMURA TETSUO
 PC C12N15/51, A61K39/29, A61K39/395, A61K39/395, C07K13/00, C12P21/02,
 C12P21/08,
 PC G01N33/53, G01N33/576, G01N33/577//A61B10/00;
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers
 FH 1. 9589
 FT source /organism='Hepatitis non-A non-B virus' FT
 FT /strain='HCV-36'.
 FEATURES
 source
 Location/Qualifiers
 1. 9589
 /organism='unidentified'
 /mol_type='genomic RNA'
 /db_xref='taxon:32644'
 BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.16% Indels: 0
 DB: 6 Gaps: 0
 US-09-980-559-2 (1-3033) x E07362 (1-9589)
 QY 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln 20
 DB 341 ATGACCAATCTTAACCTCAAGAAACCAAAAGAAACACCAACCGTCCGCCAAC 400
 QY 21 AspValLysPheProGlyGlyGlyGlnIleValGlyValTyrLeuLeuProArgArg 40
 DB 401 GACGTTAAGTTTCGGGGCGGCGGAGATCGTTGGCGAGTATACTTGTTCGCGGAGG 460
 QY 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly 60

9281 TTGCGGAGGCGGCTCTCCGATTTATCCAGCTGGTTCACTGTCGGCGCGCGGGGCG 9340
 3001 AspileTyHiserValserArgAlaargProargLeuLeuPheGlyLeuLeu 3020
 9341 GACATTTATCACAGGCTGTCGGTCCGACCGCCGCTTATCTCTTGGCTTACTCTTA 9400
 3021 LeuPheValGlyValGlyLeuPheLeuLeuProAlaArg 3033
 9401 CTTTTTGTAGGGGTAGGCTTTTCTTACTCCCGCTCGG 9439

RESULT 6
 E07361
 LOCUS 9589 bp RNA linear PAT 29-SEP-1997
 DEFINITION SRNA of Hepatitis non-A non-B virus, HCV-J6.
 ACCESSION E07361
 VERSION E07361.1 GI:5708558
 KEYWORDS JP 1994121689-A/1.
 SOURCE unidentified
 ORGANISM unidentified

REFERENCE
 1 (bases 1 to 9589)
 Okamoto, H. and Nakamura, T.
 NON-A NON-B HEPATITIS VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
 ANTIGEN AND ANTIBODY DETECTION SYSTEM
 Patent: JP 1994121689-A 1 06-MAY-1994;
 NAKAMURA TETSUO

JOURNAL
 NAKAMURA TETSUO

COMMENT
 OS Hepatitis non-A non-B virus
 PN JP 1994121689-A/1
 PD 06-MAY-1994
 PF 09-AUG-1991 JP 1991287402
 PI OKAMOTO HIROAKI, NAKAMURA TETSUO
 PC C12P21/02, A61K39/395, A61K39/395, C07K13/00, C12P21/02,
 C12P21/08,
 PC GOIN33/53, GOIN33/576, GOIN33/577//A61B10/00;
 CC strandedness: Single;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers
 FT source 1. 9589
 FT /organism="Hepatitis non-A non-B virus" FT
 /strain="HCV-J6".
 Location/Qualifiers
 1. 9589
 /organism="unidentified"
 /mol_type="genomic RNA"
 /db_xref="taxon:32644"

BASE COUNT 1968 a 2820 c 2635 g 2166 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0 Length: 9589
 Score: 15824.00 Matches: 2966
 Percent Similarity: 98.78% Conservative: 30
 Best Local Similarity: 97.79% Mismatches: 37
 Query Match: 98.16% Indels: 0
 DB: 6 Gaps: 0

US-09-980-559-2 (1-3033) x E07361 (1-9589)

Qy 1 MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgProGln 20
 Db 341 ATGAGCACAAATCTTAACCTCAAGAAAGAAACCAACCAACCGTCGCCCA 400
 Qy 21 AspValLysPheProGlyGlyGlnPheValGlyGlyValTyrlleuLeuProArg 40
 Db 401 GAGCTTAAGTTTCGGCGCGCGCGCAATGCTTGGCGGAGTACTTGTTCGCCG 460
 Qy 41 GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArg 60
 Db 461 GCGCCCGAGTTGGGTGTGCGCGCGCAAGGAAGACTTCGAGCGGTCCCGACCG 520

CDS

```
341...9442
db_xref="taxon:111103"
/codon_start=1
/product="polyprotein"
/protein_id="BAA00792.1"
/db_xref="GI:221651"
/translations="MSTNPKPKQKTRNRRPQDVKPGGQIVGVSYLLPRGRPL
GVARTKRSQPRGRQPIPKDRSTGSKSGKPGYPMPLVYNGEGLGWAGMLLSPRG
SRFSGPNRPDRNRVGVKIVITLTCGGADLMGYIPVVGAPLGGVARALAHGVRLVD
GVNFATNP,PGCSFSIFLLALISCTIPVSAAEVKNISTGVMTNDCNDISITWQIQA
AVLVHPCVPCVKRQNTSCWTPVSGNVAQVQPGALTQGLRTHIDMVMSATLCSALY
VGLD,CGVSGVLAQOMFIVTSPOHNFVQDCNSIPGTITGRNDAIMDNMSSTPATIL
```

<hr/>					
Alignment Scores:					
Pred. No.:	0	Length:	9589		
Score:	15930.00	Matches:	2967		
Percent Similarity:	98.81%	Conservative:	30		
Best Local Similarity:	97.82%	Mismatches:	36		
Query Match:	98.20%	Indels:	0		
DB:	14	Gaps:	0		
<hr/>					
US-09-980-559-2 (1-3033) x HPCPOLP (1-9589)					
Qy	1	MetSerThrAsnProLysProGlnArgLysThrLysArgAsnThrAsnArgArgProGln	20		
Dd	341	ATGAGCCAAATCCTAAACCTCAAGAAGAACCAAAAGAAAACACCACCGTCGCCCAAA	400		
..	21	AspValLysPheProGlyGlyGlyGlnIleValGlyGlyValTyLeuLeuProArgArg	40		
Qy	401	GACGTTAAGTTTCGGCGCGCGCCAGATCGTTGGCGGATATACTTCTTCGC CGCAGG	460		
Dd					
Qy	41	GlyProArgLeuGlyValArgAlaThrArgLysThrSerGluArgSerGlnProArgGly	60		
Dd	461	GGCCCCAGGTTGGTGTCGCGCGCAACAAGGAAGACTTCGAGCGGTCCACGCCACGTGGA	520		
Qy	61	ArgArgGlnProIleProLysAspArgArgSerThrGlyLysSerTrpGlyLysProGly	80		

8461 CATGTTTCACACGAGGCGAGACTGGGGTACAGCGTTGCGCGCCAGCGGGTCT 8520
 8521 CACCCTAGCATGGGAGACCATCACTGCTACGTGAAAGCTTAGCGGCTTGAAGC 8580
 8521 TACCCTAGCATGGGAGACCATCACTGCTACGTGAAAGCTTAGCGGCTTGAAGC 8580
 8581 TGCGGGATTAATCGCGCCCAATGCTGGTATGCGGCGATGACTTGGTGTCTCAGA 8640
 8581 TGCGGGATTAATGCGCGCCCAATGCTGGTATGCGGCGATGACTTGGTGTCTCAGA 8640
 8641 AAGCCAGGGGACCGAGGAGGACGAGCGGACCTGAGAGCTTCAAGGGGTATGACCG 8700
 8641 GAGCCAGGGGACCGAGGAGGACGAGCGGACCTGAGAGCTTCAAGGGGTATGACCG 8700
 8701 GTATTCGCGCTCTGCTGAGACCCCGCCAGACCGGAGTATGATCTGGAGCTGATACATC 8760
 8701 GTATTCGCGCTCTGCTGAGACCCCGCCAGACCGGAGTATGATCTGGAGCTGATACATC 8760
 8761 TTGCTCTCAATATGCTGCTGGGCGCTGGGCGCCACAGGCGCGCGAGATACCTGAC 8820
 8761 TTGCTCTCAATATGCTGCTGGGCGCTGGGCGCCACAGGCGCGCGAGATACCTGAC 8820
 8821 CAGAGACCTTACCTCCATTCGCGCGGCTGCTGGGAAACAGTTAGACACTCCCTGT 8880
 8821 CAGAGACCTTACCTCCATTCGCGCGGCTGCTGGGAAACAGTTAGACACTCCCTGT 8880
 8881 CAATTCAGGCTGGGAGACCATCATGAGTACGCGCGGACCATATGGGCTGGATGCTCT 8940
 8881 CAATTCAGGCTGGGAGACCATCATGAGTACGCGCGGACCATATGGGCTGGATGCTCT 8940
 8941 GATGACACATCTTCTCCATTCGCTGAGTGGCTCAGACACGCTGGACCAACCTCAACTT 9000
 8941 GATGACACATCTTCTCCATTCGCTGAGTGGCTCAGACACGCTGGACCAACCTCAACTT 9000
 9001 TGAGATGATCAGGAGCGGTACTCCGAGTCTGCTGGAGTCTCCAGCTATTAATGAAG 9060
 9001 TGAGATGATCAGGAGCGGTACTCCGAGTCTGCTGGAGTCTCCAGCTATTAATGAAG 9060
 9061 GTTACATGGGCTTACGCTTTCTCTGACACATACACTCCCGACGACTGACACGGGT 9120
 9061 GTTACACGGGCTTACGCTTTCTCTGACACATACACTCCCGACGACTGACACGGGT 9120
 9121 GGTTCAGCCCTCAGAAAACCTTGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGCAGC 9180
 9121 GGTTCAGCCCTCAGAAAACCTTGGGCGCCACCCCTCAGAGCGTGAAGAGCGGGCAGC 9180
 9181 TGCACTCAGGCGCTCCCTCATCTCCGCTGGGCGGAGAGCGCGGTTGCGGTGATATCT 9240
 9181 TGCACTCAGGCGCTCCCTCATCTCCGCTGGGCGGAGAGCGCGGTTGCGGTGATATCT 9240
 9241 CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCACTTCGCGGAGCGCGCTCT 9300
 9241 CTTCAATTTGGGCGGTGAAGACCAAGCTCAAACTCACTCACTTCGCGGAGCGCGCTCT 9300
 9301 GGATTTATCAGCTGGTTACCTGCGCGCGCGCGCGCGGCGGACATTTATCAAGCGTGC 9360
 9301 GGATTTATCAGCTGGTTACCTGCGCGCGCGCGCGCGGCGGACATTTATCAAGCGTGC 9360
 9361 GCGTCCCGACCCCGCTTATGCTCTTGGCGCTACTCTCTACTTTTGTAGGGGTAGGCT 9420
 9361 GCGTCCCGACCCCGCTTATGCTCTTGGCGCTACTCTCTACTTTTGTAGGGGTAGGCT 9420
 9421 TTTCTACTCTCCGCTGGTAGAGCGGACATTAAGTCACTCACTCCATTAAGTCACTGTC 9480
 9421 TTTCTACTCTCCGCTGGTAGAGCGGACATTAAGTCACTCACTCCATTAAGTCACTGTC 9480
 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 9481 CTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9540
 9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588

9541 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 9588

152 R 157 166

112861
 LOCUS 112861
 DEFINITION Sequence 2 from patent US 5428145.
 ACCESSION 112861
 VERSION 112861.1
 KEYWORDS GI:910242
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9589)
 AUTHORS Okamoto,H. and Nakamura,T.
 TITLE Non-A, non-B, hepatitis virus genome, polynucleotides, polypeptides, antigen, antibody and detection systems
 JOURNAL Patent: US 5428145-A 2 27-JUN-1995;
 FEATURES
 source Location/Qualifiers
 1. 9589
 /organism="unknown"
 BASE COUNT 1968 a 2820 c 2635 g 2166 t

Query Match 92.4%; Score 8970.4; DB 6; Length 9589;
 Best Local Similarity 96.0%; Pred. No. 0;
 Matches 9202; Conservative 0; Mismatches 386; Indels 0; Gaps 0;

QY 1 ACCCGCCCTTAATAGGGGCGACACTCCGCCATGAATCACTCCCTGTGAGGAACTACTGT 60
 DB 1 ACCCGCCCTTAATAGGGGCGACACTCCGCCATGAACCACTCCCTGTGAGGAACTACTGT 60
 QY 61 CTTCAAGCAGAAAGCGTCTAGCCATGCGGCTTAGTATGATGCTGTACAGCTCCAGGCC 120
 DB 61 CTTCAAGCAGAAAGCGTCTAGCCATGCGGCTTAGTATGATGCTGTACAGCTCCAGGCC 120
 QY 121 CCGCTCTCCGGGAGGACCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTCGCGG 180
 DB 121 CCGCTCTCCGGGAGGACCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAAATTCGCGG 180
 QY 181 AAGATGGGTCTCTTCTTGGATAAAACCACTCTATGCGCGGCCATTTGGGCGCTGCCCCCG 240
 DB 181 AAGATGGGTCTCTTCTTGGATAAAACCACTCTATGCGCGGCCATTTGGGCGCTGCCCCCG 240
 QY 241 CAAGACTGTCTAGCCAGTAGCGTTCGGTTGCGAAAGGCTTGTGGTACTGCTCTGATAGGG 300
 DB 241 CAAGACTGTCTAGCCAGTAGCGTTCGGTTGCGAAAGGCTTGTGGTACTGCTCTGATAGGG 300
 QY 301 TGCCTTGGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAAAATCTTAAACC 360
 DB 301 TGCCTTGGAGTGCCTCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAAAATCTTAAACC 360
 QY 361 TCAAGAAAAACCAAAAGAAAAACAACCGTGCGCCCAAGAGCGTTTAAAGTTTCCGGCGG 420
 DB 361 TCAAGAAAAACCAAAAGAAAAACAACCGTGCGCCCAAGAGCGTTTAAAGTTTCCGGCGG 420
 QY 421 CGCCAGATCTGTTGGCGAGTATCTTGTTCGCGCGAGGGGCCCGGAGTTGGGTGCG 480
 DB 421 CGCCAGATCTGTTGGCGAGTATCTTGTTCGCGCGAGGGGCCCGGAGTTGGGTGCG 480
 QY 481 CGCGAAGAGAGACTTCGGAGCGGTCCAGCCAGCTGCGAGCGCGCGCGCATCTCTAA 540
 DB 481 CGCGAAGAGAGACTTCGGAGCGGTCCAGCCAGCTGCGAGCGCGCGCGCATCTCTAA 540
 QY 541 AGATCGGCGCTCACTGGCAATCTGGGAAAAACCAAGGATACCTTGGCGCTTATACCG 600
 DB 541 AGATCGGCGCTCACTGGCAATCTGGGAAAAACCAAGGATACCTTGGCGCTTATACCG 600
 QY 601 GAATGAGGACTCGGCTGGGAGGATGGCTCTGCTCCCGCGAGGTTCCCGTCTCTTGG 660
 DB 601 GAATGAGGACTCGGCTGGGAGGATGGCTCTGCTCCCGCGAGGTTCCCGTCTCTTGG 660
 QY 661 GGSCCCCAATGACCCCGGCATAGTGTGCGCAACGCTGGGTAAAGTCACTGATACCTTAAC 720

